Lie Group Observer Design for Robotic Systems: Extensions, Synthesis, and Higher-Order Filtering

by

David Evan Zlotnik

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in The University of Michigan
2018

Doctoral Committee:

Adjunct Professor James R. Forbes, Chair
Associate Professor James Cutler
Associate Professor Ryan Eustice
Professor Ilya V. Kolmanovsky

David Evan Zlotnik
dzlotnik@umich
ORCID iD: 0000-0002-2995-4399

For Talya

Acknowledgements

I would first like to thank my supervisor, Prof. James R. Forbes, for providing continuous support and encouragement over the entirety of my academic career. His mentorship and guidance has been invaluable and his optimism, excitement, and passion for learning will continue to inspire me.

I would also like to thank the members of my dissertation committee, Prof. James Cutler, Prof. Ryan Eustice, and Prof. Ilya Kolmanovsky, for taking the time to review my work. I am also grateful to Prof. Timothy Barfoot for providing the experimental data used in this dissertation.

This dissertation would not be possible without the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada PGS-D program, application number PGSD3-471533-2015, and the National Science Foundation (NSF), award number 1550103. For their support I am very grateful.

I would also like to thank my fellow Ph.D. students, including Dr. Ryan Caverly and Dr. Alex Walsh, for their help and from whom I continue to learn. I would also like to thank my friends and colleagues on the Flyer team at Kitty Hawk and at Mitsubishi Electric Research Laboratory.

Finally, I would like to thank my family for their love and support, without which this dissertation would never have been written.

Preface

The contributions of this thesis are listed here in full. Modified versions of most of this dissertation have appeared as archival journal or conference publications. Citations for contributions that have appeared in refereed journals or conferences are provided.

The contributions related to the applications of gradient-based observers appearing in Part II are as follows.

- Chapter 4
  - A nonlinear observer on $SO(3)$ derived from a quaternion cost function with faster convergence properties that uses direct vector measurements [1].
  - A nonlinear observer on $SO(3)$ derived from a Gibbs cost function [2, 3].
  - A comparison of nonlinear observers on $SO(3)$ that show that the innovations are monotonically increasing with respect to attitude error [1].
  - Stability proofs that guarantee convergence of the nonlinear attitude estimators [1–3].
- Chapter 5
  - The design of a SLAM algorithm that evolves directly on the underlying Lie group [4].
  - Proof of convergence of the SLAM algorithm to a desired positively invariant set [4].
  - Application of the SLAM algorithm in simulation and experiment [4].
- Chapter 6
  - A comparison of gradient-based observer design methods to linear complementary filters [5].
  - Design of a higher-order nonlinear complementary filter for a system evolving on a Lie group [5].

  - Proof of convergence of the higher-order filter to the true state provided invariance and SPR conditions are met [5].
  - Coupling of a general disturbance observer with the higher-order filter and the proof of convergence of the coupled system [5].

- Chapter 7
  - Formulation of the Lie group observer design problem as a standard control problem.
  - A method for $\mathcal{H}_2$-optimal filter synthesis for the purposes of Lie group observer design.
  - A method for disturbance rejection via the internal model principle and an $\mathcal{H}_2$-optimal filter synthesis method for Lie group observer design.
  - The application of the higher-order filter to the attitude estimation problem [6].
  - The synthesis of an $\mathcal{H}_2$-optimal filter for the purposes of observer design on the Lie group $SO(3)$.

All figures, proofs, plots, illustrations, text, and numerical and experimental results in this dissertation were produced by David E. Zlotnik. Prof. James R. Forbes provided guidance and suggested edits throughout this dissertation. Experimental data was provided by Prof. Timothy D. Barfoot.

Table of Contents

List of Figures

Figure

List of Abbreviations

DCM	direction cosine matrix
EKF	extended Kalman filter
GBO	gradient-based observer
IMU	inertial measurement unit
LTI	linear time-invariant
LTV	linear time-varying
MEKF	multiplicative extended Kalman filter
PDF	probability density function
SLAM	simultaneous localization and mapping
SPR	strictly positive real
UAV	unmanned aerial vehicle

List of Symbols

$\langle \cdot, \cdot \rangle$	a generic inner product
$\langle \cdot, \cdot \rangle_x$	Riemannian metric at $x \in M$
$\langle \cdot, \cdot \rangle_{\mathfrak{g}}$	an inner product on $\mathfrak{g}$
M	a general manifold
∇	the gradient, or Riemannian connection
$f \circ g$	composition of maps f and g
TM	tangent bundle of M
$T_x M$	tangent space at $x \in M$
Tf	tangent map of f
$T_x f$	derivative of f at a point $x \in M$
G	a Lie group
$\mathfrak{g}$	the Lie algebra for the Lie group G
$M \to N$	a mapping from M to N
$x \mapsto f(x)$	indicates a mapping of an element $x \in M$ to a range $f(x) \in N$
$\|\| \cdot \|\|$	a generic norm
$\|\| \cdot \|\|_2$	the 2-norm of a vector or matrix
$\|\| \cdot \|\|_{\mathsf{F}}$	the Frobenius norm

$\|\| \cdot \|\|_{\mathfrak{g}}$	norm induced by an inner product on $\mathfrak{g}$
$\mathbf{A}$	a matrix, $\mathbf{A} \in \mathbb{R}^{n \times n}$ or $\mathbf{A} \in \mathbb{C}^{n \times n}$
$\mathbf{u}$	a column matrix, $\mathbf{u} \in \mathbb{R}^n$ or $\mathbf{u} \in \mathbb{C}^n$
$\mathbf{1}$	the identity matrix
$\mathbf{0}$	zero matrix
B	a basis of $\mathfrak{g}$
B_X	a basis of $T_X G$ formed by right translation by X of the elements of B
$S(\cdot)$	a mapping from $\mathbb{R}^n$ to $\mathfrak{g}$
$L_g(\cdot)$	left translation by $g \in G$
$R_g(\cdot)$	right translation by $g \in G$
Ad	the adjoint map
$[\cdot, \cdot]$	the Lie bracket

Abstract

The kinematics and dynamics of many robotic systems evolve on differential manifolds, rather than strictly in Euclidean space. Lie groups, a class of differential manifold with a group structure, arise naturally in the study of rigid-body kinematics. This dissertation studies the design of state observers for systems whose state evolves on a Lie group. State observers, or state estimators, are a crucial part of the guidance, navigation, and control algorithms necessary for autonomous operation of many ground, air, and marine vehicles. The design of state observers on Lie groups is therefore a highly practical exercise. One such nonlinear observer, the gradient-based observer, has generated significant interest in the literature due to its computational simplicity and stability guarantees. The first part of this dissertation explores several applications of the gradient-based observer, including both the attitude estimation problem and the simultaneous localization and mapping (SLAM) problem. By modifying the cost function associated with the observer, several novel attitude estimators are introduced that provide faster convergence when the initial attitude error is large. Further, a SLAM algorithm with guaranteed convergence is introduced and tested in both simulation and experiment.

In the second part of this dissertation, the state of the art in Lie group observer design is extended by the development of a higher-order filter on a Lie group. By analogy to the classical linear complementary filter, the proposed method can be interpreted as a nonlinear complementary filter on a Lie group. A disturbance observer that accounts for constant and harmonic disturbances in the group velocity measurements is also considered. Local asymptotic stability about the desired equilibrium point is demonstrated. In addition, an $\mathcal{H}_2$-optimal filter synthesis method is derived and disturbance rejection via the internal model principle is considered. A numerical example that demonstrates the desirable properties of the higher-order nonlinear complementary filter, as well as the synthesis techniques, is presented in the context of rigid-body attitude estimation.

Part I

Introduction

Chapter 1

Introduction

1.1 Motivation and Objectives

Guidance, navigation, and control of robotic systems can be summarized by three questions. *Where am I? Where am I going? How do I get there?* The first question is the problem of navigation, the process of estimating the state of a robotic system, such as its position, velocity, and orientation. The second question is the problem of guidance, which seeks to determine the path that a robotic system must take. The third question is the problem of control, where inputs and commands are sought such that the robotic system follows a desired trajectory. All three processes are intertwined in practice. For example, a motion planning algorithm in an autonomous vehicle determines a desired path for the vehicle, while the control system ensures the vehicle remains on the path using information from the navigation algorithm.

This dissertation is concerned with navigation. Navigation is typically accomplished by using a state observer, or state estimator, that fuses sensor data and control commands to determine the state of the robotic system. State estimation for systems whose state evolves in Euclidean space is a mature field whose history[1] begins with the introduction of the famous Kalman filter in the 1960s [7]. The Kalman filter and its nonlinear extensions, such as the extended Kalman filter, have been indispensable navigation tools that have seen use from Project Apollo, NASA's manned lunar landing missions, to the autonomous vehicle "Stanley", the winner of the 2005 DARPA Grand Challenge [8]. The Kalman filter belongs to a class of estimation strategies known as recursive Bayesian filtering [9]. The filters are recursive in that

[1]Modern state estimation concepts were formalized by Rudolf Kalman and others in the 1960s. However, navigation solutions appear throughout history, perhaps beginning with the appearance of the first seafaring boat.

current state estimate is calculated only from the previous state estimate and current measurements. This is opposed to batch estimation techniques that utilize the full time history of the state estimates and measurements. The title Bayesian refers to the fact that the filters are derived from, or are approximations of, the Bayes filter, a general state estimation technique that seeks to determine the probability density function (PDF) that represents the likelihood of the state. Consequently, recursive Bayesian filters are inherently probabilistic. Recursive Bayesian filters can be subdivided into Gaussian filters and nonparameteric filters. Gaussian filters, the group to which the Kalman filter belongs, make the assumption that the PDFs associated with the state and measurement noise are normally distributed. Alternatively, nonparametric filters, such as the particle filter, make no assumption on the form of the probability distributions [9].

The introduction of new technologies in recent years has lead to the proliferation of small and low-cost robotic systems, such as miniature unmanned aerial vehicles (UAVs). With the advent of these systems comes new and exciting estimation challenges. Low cost inertial measurement units (IMUs) often found on small UAVs, for example, are plagued by high levels of non-Gaussian noise and time-varying biases. In these instances the application of traditional probabilistic estimation methods can be difficult. Compounding difficulties is the fact that these estimation algorithms must often be implemented on embedded computers that have limited computational resources [10]. Consequently, the development of robust and computationally efficient estimators has been an active area of research over the last decade. In particular, a class of nonlinear observers that evolve on Lie groups has recently garnered significant attention [11–16]. Lie groups are a well known class of manifold that occur naturally in the study of rigid-body kinematics. Attitude kinematics, for example, evolve on the special orthogonal group $SO(3)$, while pose kinematics evolve on the special Euclidean group $SE(3)$ [17]. The development of observers for systems whose state evolves on a Lie group is therefore a highly practical exercise. The interest in these nonlinear observers was, in part, sparked by the development of nonlinear observers for attitude estimation using first the unit quaternion [18–21] and later the rotation matrix element of $SO(3)$ directly [10, 22]. Following [10], several nonlinear attitude observers that exploit the underlying $SO(3)$ Lie group structure have been developed [23–25], as well as several $SE(3)$ based nonlinear observers [26, 27]. By working directly with the elements of $SO(3)$ and $SE(3)$, the attitude and pose are both globally and uniquely represented and thus issues associated with attitude parameterizations, such as non-uniqueness, are avoided [28]. Nonlinear observers are

attractive as they can often be shown to have strong stability properties [14, 15, 21] and are, in general, computationally simpler than traditional estimation methods [14]. One such nonlinear observer, the gradient-based observer considered in [14], is of principal importance to this dissertation.

The goal of this dissertation is to extend and build upon the nonlinear observer design methods previously developed in the literature. The content of this dissertation is contained in two parts. The first part is concerned with the development of novel state estimation techniques by the application of the gradient-based observer, a Lie group observer developed in [14]. The second part extends the state of the art in general Lie group observer design by the development of a higher-order filter on a Lie group.

1.1.1 Applications of the Gradient-Based Observer

In the first part, two applications of gradient-based observer design on Lie groups are considered, the attitude estimation problem and the simultaneous localization and mapping (SLAM) problem. Gradient-based design methodology involves propagating the state estimate along the gradient descent direction of a selected error function. The behavior of the observer is therefore inherently linked to the selection of the cost function. Previous gradient-based observers for the attitude estimation problem suffer from slow convergence, especially when the initial attitude error is large, the cause of which can be traced to the initial selection of a cost function on the group $SO(3)$. The first objective is then to explore alternate attitude cost functions that result in faster convergence of the attitude estimate.

Simultaneous localization and mapping is the process of building a map of the environment surrounding a vehicle while simultaneously estimating the pose of the vehicle relative to the map. The most popular SLAM methods, such as EKF SLAM, FastSLAM, GraphSLAM, iSAM, and iSAM2 [9, 29–32], are probabilistic and provide estimates of the state as well as its associated PDF. SLAM solutions, however, rarely have guaranteed convergence properties. By formulating the SLAM problem as a gradient-based observer design problem, it will be possible to derive a computationally simple and provably convergent SLAM algorithm. The second objective is to develop and characterize a gradient-based observer for the purposes of SLAM.

1.1.2 Higher-Order Nonlinear Complementary Filtering

Complementary filtering is a simple linear estimation technique used to fuse two measurements of the same state by exploiting the complementary noise characteristics of the two measurements. The gradient-based observer can be interpreted as a nonlinear complementary filter on a Lie group. In this interpretation, current methods, including [14] and [16], are analogous to the linear complementary filter with first order sensitivity and complementary sensitivity transfer functions. The first order nature of these observers are restrictive and place limits on the their effective bandwidths and roll-off rates. Consequently, a Lie group observer that allows for higher-order filtering is highly desirable.

While the gain of the innovation terms in [14] are static, a Lie group observer with higher-order filtering would introduce dynamics in the innovation term. This is similar to dynamic observers for systems evolving on vector spaces [33–37]. Dynamic observers are frequently used to approach the problem of robust state estimation in the $\mathcal{H}_\infty$ framework, see for example in [34]. Although dynamic observers have received little attention in the literature, the problem of dynamic observer design is the dual problem of the well known and understood dynamic controller design problem. Dynamic controller design, for example $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controller design, has received significant attention over the last several decades. Consequently, the formulation of a higher-order filter on a Lie group will allow for the use of the immensely successful $\mathcal{H}_2$- and $\mathcal{H}_\infty$-optimal control strategies in the context of Lie group observer design. Thus, the main objectives of the second part of this dissertation are to (1) fully extend the concept of linear complementary filtering to higher-order filtering on Lie groups, and (2) to design an $\mathcal{H}_2$-optimal synthesis method for nonlinear observer design on Lie groups.

1.2 Outline and Contributions

In Chapter 2, preliminaries and general concepts used throughout this dissertation are reviewed including linear algebra and differential geometry. Part II focuses on applications of the gradient-based observer. An introduction to gradient-based observer design is given in Chapter 3. The two main contributions of Part II are given in Chapters 4 and 5. Attitude estimation is the focus of Chapter 4. The contribution of Chapter 4 is the design of two novel attitude estimation algorithms that result in faster convergence of the attitude estimate compared to previously developed gradient-based observers. Simultaneous localization and mapping is discussed

in Chapter 5, where a novel SLAM algorithm, with provable convergence properties, is developed.

Part III of this dissertation is dedicated to the development and characterization of a higher-order nonlinear complementary filter on a Lie group. Within Part III, Chapter 6 discusses the design of a higher-order nonlinear complementary filter and discusses the relationship between linear complementary filtering and gradient-based observer design. Stability of the proposed method is also discussed in Chapter 6. Chapter 7 is dedicated to the problem of $\mathcal{H}_2$-optimal filter synthesis. The design of the higher-order complementary filter and the $\mathcal{H}_2$-optimal synthesis method are the main contributions of Part III. The findings of this dissertation are summarized in Chapter 8, and suggested future research directions are suggested. A detailed list of contributions and publications can be found in the Preface.

Chapter 2

Preliminaries

This dissertation deals heavily with the concepts of systems and differential geometry. The concepts of signals and systems are reviewed in Sec. 2.1 while relevant concepts in differential geometry are covered in Sec. 2.2.

2.1 Signals and Systems

2.1.1 Vector Spaces and Norms

Definition 2.1 (Vector Space [17])**.** A vector space over the field $\mathbb{F}$ is a set $\mathcal{V}$ equipped with two operations: (1) vector addition, denoted by $v_1 + v_2 \in \mathcal{V}$ for $v_1, v_2 \in \mathcal{V}$, and (2) scalar multiplication, denoted by $av \in \mathcal{V}$ for $a \in \mathbb{F}$ and $v \in \mathcal{V}$. Vector addition must satisfy the rules

(i) $v_1 + v_2 = v_2 + v_1$, $v_1, v_2 \in \mathcal{V}$,

(ii) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$, $v_1, v_2, v_3 \in \mathcal{V}$,

(iii) there exists a unique vector $0 \in \mathcal{V}$ with the property that $v + 0 = v$ for every $v \in \mathcal{V}$,

(iv) for every $v \in \mathcal{V}$, there exists a unique vector $-v \in \mathcal{V}$ such that $v + (-v) = 0$,

and scalar multiplication must satisfy the rules

(i) $a_1(a_2 v) = (a_1 a_2)v$, $a_1, a_2 \in \mathbb{F}$, $v \in \mathcal{V}$,

(ii) $1v = v$, $v \in \mathcal{V}$,

(iii) $a(v_1 + v_2) = av_1 + av_2$, $a \in \mathbb{F}$, $v_1, v_2 \in \mathcal{V}$, and

(iv) $(a_1 + a_2)v = a_1 v + a_2 v$, $a_1, a_2 \in \mathbb{F}$, $v \in \mathcal{V}$.

Definition 2.2 (Isomorphic Vector Spaces [38])**.** Consider two vector fields $\mathcal{V}$ and $\mathcal{W}$. Then, $\mathcal{V}$ is isomorphic to $\mathcal{W}$, written $\mathcal{V} \cong \mathcal{W}$, if and only if there exists a linear transformation from $\mathcal{V}$ to $\mathcal{W}$, $L : \mathcal{V} \to \mathcal{W}$, such that L is one-to-one and onto.

Definition 2.3 (Linear Map [39])**.** A mapping $A : \mathcal{V} \to \mathcal{W}$ is linear if it satisfies $A(av + bu) = aA(v) + bA(u)$ for all $v, u \in \mathcal{V}$ and $a, b \in \mathbb{F}$.

Definition 2.4 (Vector Norm [39])**.** A norm $||\cdot||$ on a vector space $\mathcal{V}$ is a function that maps $\mathcal{V} \to [0, \infty)$ which for all $v \in \mathcal{V}$ that satisfies

(i) $||v|| = 0$ if an only if $v = 0$,

(ii) $||av|| = |a|\ ||v||$ for all $a \in \mathbb{F}$, and

(iii) the triangle inequality $||u + v|| \leq ||u|| + ||v||$, for all $v, u \in \mathcal{V}$.

2.1.2 Inner Product Spaces

Definition 2.5 (Inner Product [39])**.** An inner product $\langle \cdot, \cdot \rangle$ on a vector space $\mathcal{V}$ is a function mapping $\mathcal{V} \times \mathcal{V} \to \mathbb{F}$ such that

(i) for all $v \in \mathcal{V}$, $\langle v, v \rangle \geq 0$,

(ii) for all $v \in \mathcal{V}$, $\langle v, v \rangle = 0$ if and only if $v = 0$,

(iii) for all $u,\ v,\ w \in \mathcal{V}$ and $a,\ b \in \mathbb{F}$, $\langle u, av + bw \rangle = a\langle u, v \rangle + b\langle u, w \rangle$.

A vector space equipped with an inner product is called a an inner product space. An important inner product space is the space $\mathcal{L}_2$, the Lebesgue space, defined by

$$\mathcal{L}_2 = \left\{ \mathbf{x} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n \;\middle|\; \int_0^\infty \mathbf{x}(t)^\mathsf{T} \mathbf{x}(t)\ \mathrm{d}t < \infty \right\}. \tag{2.1}$$

The extended Lebesgue space, denoted $\mathcal{L}_{2e}$ is defined as

$$\mathcal{L}_{2e} = \left\{ \mathbf{x} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n \;\middle|\; \int_0^T \mathbf{x}(t)^\mathsf{T} \mathbf{x}(t)\ \mathrm{d}t < \infty,\ \forall T \in \mathbb{R}_{>0} \right\}. \tag{2.2}$$

2.1.3 Linear Systems

Definition 2.6 (Linear Time-Invariant System)**.** A system $\boldsymbol{\mathcal{H}}$ is linear time-invariant (LTI) if it is linear and if

$$\mathbf{y}(t+\tau) = \boldsymbol{\mathcal{H}}\mathbf{u}(t+\tau),\ \forall t, \tau \in \mathbb{R}.$$

An LTI system can be represented in state-space form by the system of ordinary differential equations given by

$$\boldsymbol{\mathcal{H}} \left\{ \begin{array}{rl} \dot{\mathbf{x}} & = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} & = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \end{array} \right.$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$, and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are appropriately dimensioned real matrices. The transfer matrix associated with $\boldsymbol{\mathcal{H}}$ is denoted $\mathbf{H}(s) \in \mathbb{C}^{p\times m}$ and is given by $\mathbf{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$.

Definition 2.7 ($\mathcal{H}_2$ Norm of a Linear System [40])**.** The $\mathcal{H}_2$ norm of the asymptotically stable linear system $\boldsymbol{\mathcal{H}} : \mathcal{L}_{2e} \to \mathcal{L}_{2e}$, denoted as $\|\boldsymbol{\mathcal{H}}\|_2$, is given by

$$\|\boldsymbol{\mathcal{H}}\|_2 = \sqrt{\int_0^\infty \operatorname{tr}\left(\mathbf{H}^{\mathsf{T}}(t)\mathbf{H}(t)\right) \mathrm{d}t},$$

where $\mathbf{H}(t)$ is the impulse response of the transfer matrix $\mathbf{H}(s)$. Equivalently,

$$\|\boldsymbol{\mathcal{H}}\|_2 = \sqrt{\frac{1}{2\pi}\int_{-\infty}^\infty \operatorname{tr}\left(\mathbf{H}^{\mathsf{H}}(j\omega)\mathbf{H}(j\omega)\right) \mathrm{d}\omega}.$$

2.2 Differential Geometry

This dissertation is primarily concerned with the development of observers for systems whose state evolves not necessarily in Euclidean space but on more complex topological spaces. Consequently, it is necessary to venture into the field of differential geometry, which allows for the study of dynamical systems on complex geometric structures. More generally, differential geometry is the study of geometric problems through the lens of calculus and linear algebra.

2.2.1 Differential Manifolds and the Tangent Space

The fundamental structure of differential geometry is the differential manifold, which is defined below.

Definition 2.8 (Differential Manifold [41])**.** An n-dimensional differential manifold M is a set of points together with a finite or countably infinite set of subsets $U_\alpha \subset M$ and one-to-one mappings $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ such that:

(i) $\cup_\alpha U_\alpha = M$;

(ii) for each nonempty intersection $U_\alpha \cap U_\beta$, $\phi_i(U_\alpha \cap U_\beta)$ is an open subset of $\mathbb{R}^n$, and the one-to-one and onto mapping $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$ is a smooth function;

(iii) the family $\{U_\alpha, \phi_\alpha\}$ is maximal with respect to conditions (i) and (ii).

The definition of a differential manifold presented in Definition 2.8 is quite abstract. Less formally, a differential manifold may be understood as a topological space that is locally indistinguishable from Euclidean space. A curve on a manifold M is defined as a smooth function $\gamma : \mathbb{R} \to M : t \mapsto \gamma(t)$ [42]. As pointed out in [41], it should be possible to take the derivative of $\gamma(\cdot)$ as M is locally indistinguishable from Euclidean space and differentiation involves only local computations. However, defining the derivative in the classical sense as $\gamma'(t) \triangleq \lim_{h\to 0}(\gamma(t+h) - \gamma(t))/h$ is not possible for any manifold as the computation of the difference $\gamma(t+h) - \gamma(t)$ implies some vector space structure of the manifold itself [42]. Consequently, the definition of the derivative of a map on a manifold must be more carefully considered. To start consider first a curve on a manifold.

Definition 2.9 (Curve at a Point [17])**.** Let $x \in M$. A curve at x is a mapping $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an open interval containing 0 in its interior, and for which $\gamma(0) = x$.

Definition 2.10 (Equivalent Curves [17, 41])**.** Consider a point $x \in M$ and let γ_1 and γ_2 be curves at x. The curves γ_1 and γ_2 are equivalent at x, written $\gamma_1 \sim_x \gamma_2$, if for all real valued functions $f : M \to \mathbb{R}$ defined on an open neighborhood of x the relation

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_1(t))\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_2(t))\bigg|_{t=0} \tag{2.3}$$

holds. The equivalence class of a curve γ at x, denoted $[\gamma]_x$, is the set of curves $[\gamma]_x = \{\gamma_1 : I \to M \mid \gamma_1 \sim_x \gamma\}$.

An equivalence class of curves can be thought of as the combination of all curves that pass through x in the same direction. The equivalence class can be said to describe the direction itself, and a single equivalence class of curves is called a tangent vector.

Definition 2.11 (Tangent Vector [17])**.** A tangent vector is an equivalent class of curves under the equivalence relation $\sim_x$.

Combining the set of all equivalence classes of curves at a point x forms a vector space called the tangent space at x. Further combining all of the tangent spaces on a manifold forms another manifold called the tangent bundle.

Definition 2.12 (Tangent Space [17, 41])**.** The tangent space at a point $x \in M$, denoted T_xM, is the vector space formed by the collection of all tangent vectors at x.

Definition 2.13 (Tangent Bundle [41])**.** The tangent bundle of M, denoted TM, is the differential manifold whose underlying set is the disjoint union of the tangent spaces at the points $x \in M$; that is, $TM = \cup_{x \in M} T_xM$.

The above definitions of a tangent vector and tangent space are, again, quite abstract. This is necessitated by the fact that the manifold under consideration is a general manifold. There are in fact several alternative definitions of the tangent space found in the literature. For example, in [43] the tangent space is defined as the vector space formed from the set of all derivations at a point on the manifold. These definitions are all equivalent in the sense that the different definitions of the tangent space are all isomorphic to each other. What is important to note is that the tangent space at x is a vector space that encodes all possible directions of a curve passing through x. It is now possible to discuss the concept of the derivative of a map between two manifolds.

Definition 2.14 (Derivative of a Function [17, 41])**.** Let $f : M \to N$ be a mapping between differential manifolds M and N. The derivative of f at a point $x \in M$ is the linear map $T_xf : T_xM \to T_{f(x)}N$ defined as $T_xf([\gamma]_x) = [f \circ \gamma]_{f(x)}$ for all $[\gamma]_x \in T_xM$.

The map $Tf : TM \to TN$ is referred to as the tangent map of f. To illustrate Definition 2.14, suppose γ is a curve at a point $x \in M$ and let $f : M \to N$. Then, $[\gamma]_x$, an equivalence class of curves at x, defines a tangent vector on T_xM. The derivative of f at x is a mapping that maps tangent vectors in T_xM to $T_{f(x)}N$. When T_xf is evaluated at $[\gamma]_x$, the resulting tangent vector on $T_{f(x)}N$ is given by the equivalence class $[f \circ \gamma]_{f(x)}$. The mapping $T_xf(v)$, with $v \in T_xM$, can be computed by first

defining a curve at x, $\gamma(\cdot)$, such that $\gamma(0) = x$ and $\mathrm{d}\gamma/\mathrm{d}t|_{t=0} = v$. Then, $T_x f(v)$ is given by [41]

$$T_x f(v) = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\bigg|_{t=0}. \tag{2.4}$$

When the mapping is a curve $\gamma : I \to M$, $t \mapsto \gamma(t)$, the notation $T\gamma(\cdot) = \dot{\gamma}(\cdot)$ will be adopted, where $\dot{\gamma}(t) \in T_{\gamma(t)}M$.

The flow of a curve on a manifold can be described by a vector field, defined below.

Definition 2.15 (Vector Field [41])**.** A vector field, V, on M is a map $V : M \to TM$ that assigns a tangent vector $V(x) \in T_x M$ for every point $x \in M$.

2.2.2 Lie Groups and Lie Algebras

An important type of manifold that will be discussed in detail in this dissertation is the Lie group defined as follows.

Definition 2.16 (Lie Group [41])**.** A Lie group is a smooth manifold, denoted G, that is a group and for which the group operations

(i) $(g, h) \mapsto gh$, for $g, h \in G$, and

(ii) $g \mapsto g^{-1}$, for $g \in G$,

are smooth.

Definition 2.17 (Left and Right Translation Maps [17])**.** The left translation map is the map $L_g : G \to G$, $h \mapsto hg$, for all $h, g \in G$. Equivalently, the right translation map is the map $R_g : G \to G$, $h \mapsto gh$.

There exists an isomorphism between the tangent space at the identity element of G and the tangent space at any point $g \in G$. This is written $T_e G \cong T_g G$ [17]. Thus tangent vectors at a point $g \in G$ can be identified with elements of $T_e G$ by either the map $T_e L_g : T_e G \to T_g G$, $\xi \mapsto T_e L_g(\xi)$, or $T_e R_g : T_e G \to T_g G$, $\xi \mapsto T_e R_g(\xi)$. For convenience, the tangent vectors $T_e L_g(\xi)$ and $T_e R_g(\xi)$ will often be abbreviated as $g\xi$ and ξg respectively [17, 41]. When g varies over the group the vectors $g\xi$ and ξg define left- and right-invariant vector fields respectively [41]. The left-invariant vector field generated by ξ is denoted $\xi_L : G \to TG$ defined by $\xi_L(g) = g\xi$. A vector field, $V : G \to TG$, is said to be left-invariant if $V(gh) = T_h L_g(V(h))$. Similarly, a vector field is right-invariant if $V(hg) = T_h R_g(V(h))$ [17].

Definition 2.18 (Lie Algebra [17])**.** A Lie algebra is a vector space, $\mathcal{V}$, endowed with a bilinear operation $[\cdot,\cdot]:\mathcal{V}\times\mathcal{V}\to\mathcal{V}$ satisfying

(i) $[a,b]=-[a,b]$ for all $a,b\in\mathcal{V}$, and

(ii) $[a,[b,c]]+[b,[c,b]]+[c,[a,b]]=0$ for all $a,b,c\in\mathcal{V}$.

Definition 2.19 (Lie Algebra of a Lie Group [17])**.** The Lie algebra of a Lie group G, denoted $\mathfrak{g}$, is the tangent space at the identity element $e\in G$, $\mathfrak{g}=T_eG$, with the corresponding Lie bracket $[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ given by

$$[\xi,\eta]=[\xi_L,\eta_L](e), \tag{2.5}$$

where $[\xi_L,\eta_L]$ is the Jacobi-Lie bracket of vector fields ξ_L and η_L.

Definition 2.20 (Adjoint Map [17])**.** The adjoint map, denoted $\mathrm{Ad}_X:\mathfrak{g}\to\mathfrak{g}$ is a linear map defined by

$$\mathrm{Ad}_X(v)=T_XR_{X^{-1}}(T_IL_X(v)), \tag{2.6}$$

for all $X\in G$ and $v\in T_IG$.

The derivative of the adjoint map, $T\mathrm{Ad}_X$, at a point $v\in\mathfrak{g}$, is denoted ad_v, where it can be shown that $\mathrm{ad}_v(u)=[v,u]$ for all $v,u\in\mathfrak{g}$ [44].

2.2.3 Matrix Lie Groups and Matrix Lie Algebras

In this dissertation the Lie group under consideration will often be a matrix Lie group. Matrix Lie groups arise in the discussion of rigid-body kinematics and dynamics. For example, the attitude of a rigid body can be described by the matrix Lie group $SO(3)$, which is the set of all 3×3 orthonormal matrices with determinant equal to $+1$. A matrix Lie group is defined formally below.

Definition 2.21 (Matrix Lie Group [41])**.** A matrix Lie group is a subgroup of the general linear group, denoted $GL(n,\mathbb{R})$, which is the set of invertible $n\times n$ real matrices.

Let $\mathbf{X}$ be an element of a matrix Lie group G. From Definition 2.12, the tangent space at $\mathbf{X}$, $T_{\mathbf{X}}G$, is defined as the set of all equivalent classes of curves passing through $\mathbf{X}$. However, matrix Lie groups are in fact embedded submanifolds of $\mathbb{R}^{n\times n}$ [42]. Due to the vector space structure of $\mathbb{R}^{n\times n}$ the time derivative of a curve $\mathbf{\Gamma}:I\to G$ on a matrix

Lie group G at $\mathbf{X} \in G$ is well defined and is given by $\dot{\mathbf{\Gamma}}(t) = \lim_{h\to 0}(\mathbf{\Gamma}(t+h)-\mathbf{\Gamma}(t))/h$. Consequently, alternative definitions of the tangent space and Lie algebra of a matrix Lie group as linear subspaces of $\mathbb{R}^{n\times n}$ may be defined as follows.

Definition 2.22 (Tangent Space of a Matrix Lie Group [42])**.** The tangent space of a matrix Lie group G at a point $\mathbf{X} \in G$ is the vector space given by $T^{\text{mat}}_{\mathbf{X}}G = \{\dot{\mathbf{\Gamma}}(0) \in \mathbb{R}^{n\times n} \mid \mathbf{\Gamma} \text{ a curve at } \mathbf{X}\}$.

Definition 2.23 (Lie Algebra of a Matrix Lie Group [17, 41])**.** The Lie algebra of a matrix Lie group G, is the tangent space at the identity element $\mathbf{1} \in G$, $\mathfrak{g}^{\text{mat}} = T^{\text{mat}}_{\mathbf{1}}G$, with the corresponding Lie bracket $[\cdot,\cdot] : \mathfrak{g}^{\text{mat}} \times \mathfrak{g}^{\text{mat}} \to \mathfrak{g}^{\text{mat}}$ given by $[\mathbf{A},\mathbf{B}] = \mathbf{AB} - \mathbf{BA}$.

The superscript "mat" is used to differentiate the above definitions of tangent space and Lie algebra with the definitions given in Definition 2.12 and Definition 2.19. However, these definitions are equivalent in the sense that $T^{\text{mat}}_{\mathbf{X}}G \cong T_{\mathbf{X}}G$ and $\mathfrak{g}^{\text{mat}} \cong \mathfrak{g}$. Thus, any element of $T_{\mathbf{X}}G$ or $\mathfrak{g}$ may be identified by an element of $T^{\text{mat}}_{\mathbf{X}}G$ or $\mathfrak{g}^{\text{mat}}$, respectively. Therefore, the superscript will be dropped in the remainder of this dissertation and it will understood from context which definitions are being used.

2.2.4 Riemannian Metric, Gradient, and Hessian

Consider a differential manifold M and let $x \in M$. An inner product, denoted $\langle\cdot,\cdot\rangle_x : T_xM \times T_xM \to \mathbb{R}$, may be defined for each tangent space T_xM. When this inner product is smoothly varying the inner product is referred to as a Riemannian metric and the manifold as a Riemannian manifold [42]. The Riemannian metric allows for the definition of a gradient of a function on a manifold as follows.

Definition 2.24 (Gradient [41, 42])**.** The gradient of a function $f : M \to \mathbb{R}$, denoted $\nabla f(x) \in T_xM$, is the unique tangent vector satisfying

$$T_xf(v) = \langle \nabla f(x), v\rangle_x, \ \forall v \in T_xM. \tag{2.7}$$

The mapping $\nabla f : M \to TM$ is often called the gradient vector field [41]. When f is a mapping $f : G \times G \to \mathbb{R}$, the notation $\nabla_x f(x,y)$, $x, y \in G$, will denote the gradient of f with respect to x. Associated with every Riemannian metric is a unique torsion free and affine connection ∇ called the Riemannian connection, that assigns to each pair of vector field V and U a vector field $\nabla_V U$ [42].

Definition 2.25 (Affine Connection [41, 42])**.** An affine connection ∇ on a manifold M is a mapping that assigns to each pair of vector fields U and V a vector field $(U, V) \mapsto \nabla_U V$ that satisfies the following properties:

(i) $\nabla_{fU+gV} W = f\nabla_U W + g\nabla_V W$;

(ii) $\nabla_U (V + W) = \nabla_U V + \nabla_U W$;

(iii) $\nabla_U (fV) = f\nabla_U V + (\mathscr{L}_U f)V$;

for all smooth functions $f : M \to \mathbb{R}$ and $g : M \to \mathbb{R}$ and vector fields U, V, and W, where $fV : M \to TM$, $x \mapsto f(x)V(x)$.

Definition 2.26 (Riemannian Connection [41, 42])**.** The Riemannian connection, also called the Levi-Civita connection, of a Riemannian manifold is the unique affine connection ∇ such that

(i) $\nabla_U V - \nabla_V U = [U, V]$, and

(ii) $W\langle U, V\rangle = \langle \nabla_W U, V\rangle + \langle U, \nabla_W V\rangle$.

Definition 2.27 (Riemannian Hessian [42])**.** The Riemannian Hessian, or Hessian operator, of a function $f : M \to \mathbb{R}$ at x is the linear mapping $H_{f(x)} : T_x M \to T_x M$ defined by

$$H_{f(x)}(V(x)) = \nabla_V \nabla f(x) \tag{2.8}$$

Part II

Applications of the Gradient-Based Observer

Chapter 3

Introduction to Gradient-Based Observer Design

3.1 Introduction

In this chapter, the gradient-based observer design methodology for invariant systems on Lie groups developed in [14] is reviewed. Gradient-based observers are a class of nonlinear observers whose innovation term is derived from the gradient of a chosen cost function. The results of [14], and its extensions in [15, 16] are quite remarkable, in that the observer design methodology results in computationally simple nonlinear observers with strong stability guarantees. Moreover, the observer design methods are applicable to a large variety of different robotic systems, from autonomous aerial vehicles to robotic manipulators, due to the fact that the observer design problem is posed for any general Lie group with invariant dynamics.

The gradient-based observer on a Lie group can be described as follows. First, a copy of the system dynamics is used to propagate the state along the trajectory of the true system. An innovation term, constructed from the gradient of a cost function, is added to ensure that the state estimate approaches the true state in the presence of initialization or measurement errors. Invariance of the cost function, as well as the selection of an invariant Riemannian metric on the Lie group, ensures that the error dynamics associated with the observer is autonomous. This again is quite remarkable as it greatly simplifies the stability analysis of the proposed observers. The gradient-based observer proposed in [14] is extended in [16] where the innovations are based on gain mappings applied to the differential, rather than the gradient, of a cost function.

Gradient-based observer design methodology is of significant importance to this

dissertation. The design methodology is used to develop novel nonlinear attitude estimators in Chapter 4 and is applied to the problem of simultaneous localization and mapping in Chapter 5. Further, gradient-based observer design is extended to the concept of higher-order filtering in Part II of this dissertation.

3.2 System and System Outputs

3.2.1 System Equations

Let G denote a finite-dimensional Lie group and consider a curve $X : \mathbb{R} \to G$ on G governed by the differential equation

$$\dot{X}(t) = V(t, X(t)), \tag{3.1}$$

where $V : \mathbb{R} \times G \to TG$ is a time-dependent vector field and $\dot{X}(t) \in T_{X(t)}G$ is a tangent vector at $X(t)$. For the purposes of this dissertation it will will be convenient to express $V(t, X(t))$ as the left-invariant vector field $V(t, X(t)) = X(t)v(t)$, where $v : \mathbb{R} \to \mathfrak{g}$ is an exogenous signal. The quantity $v(\cdot)$ is often called the group velocity. Doing so, (3.1) can be expressed as

$$\dot{X}(t) = X(t)v(t). \tag{3.2}$$

3.2.2 Measurements

It is assumed that a measurement of $v(\cdot)$ is available as

$$v_y(t) \;\; = \;\; v(t) + d(t), \tag{3.3}$$

where $v_y \in \mathfrak{g}$ is the measurement of v, and $d \in \mathfrak{g}$ is the noise associated with measurement v_y.

It is also assumed that $\ell \in \mathbb{N}$, $\ell > 0$, partial state measurements of X are available as

$$y_j(t) = h_j(N_j(t)X(t), \bar{y}_j), \;\; j \in \{1, \ldots, \ell\}, \tag{3.4}$$

where $N_j \in G$ is multiplicative noise associated with the measurement of y_j. As in [45], the partial state measurements, or system outputs, y_j are assumed to be elements of a manifold $\mathcal{M}$ and $\bar{y}_j \in \mathcal{M}$ are constant reference outputs. Multiplicative noise, $N_j \in G$, rather than additive noise on $\mathcal{M}$ is chosen as $\mathcal{M}$ is not necessarily a vector space. The mappings $h_j : G \times \mathcal{M} \to \mathcal{M}$ are assumed to be right actions of G on $\mathcal{M}$

such that $h_j(I, y) = y$ and $h_j(X, h_j(Y, y)) = h_j(XY, y)$ for all $X, Y \in G$ and $y \in \mathcal{M}$ [16, 45]. In the remainder of this chapter, the argument of time, t, will be dropped for time dependent variables when appropriate.

For simplicity of notation, define $\mathcal{M}^\ell = \mathcal{M} \times \cdots \times \mathcal{M}$ and $G^\ell = G \times \cdots \times G$. Then, letting

$$y = (y_1, \ldots, y_\ell) \in \mathcal{M}^\ell, \tag{3.5}$$

$$\bar{y} = (\bar{y}_1, \ldots, \bar{y}_\ell) \in \mathcal{M}^\ell, \tag{3.6}$$

$$N = (N_1, \ldots, N_\ell) \in G^\ell, \tag{3.7}$$

the system outputs may be written as

$$y = h(X, N, \bar{y}), \tag{3.8}$$

where $h : G \times G^\ell \times \mathcal{M}^\ell \to \mathcal{M}^\ell$, $(X, N, \bar{y}) \mapsto (h_1(N_1 X, \bar{y}_1), \ldots, h_\ell(N_\ell X, \bar{y}_\ell))$.

3.3 Gradient-Based Observer Design

3.3.1 Linear Observer Design

Traditional observer design for systems evolving in Euclidean space, such as the Luenberger observer, can be described by the propagation of the state estimate using a process model and a correction term. For example, consider the linear time-invariant system and associated Luenberger observer

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{3.9}$$

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}\mathbf{C}(\mathbf{x} - \hat{\mathbf{x}}), \tag{3.10}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the system state, $\mathbf{y} = \mathbf{C}\mathbf{x} \in \mathbb{R}^p$ is the system output, $\mathbf{u} \in \mathbb{R}^m$ is the system input, $\hat{\mathbf{x}} \in \mathbb{R}^n$ is the state estimate, $\hat{\mathbf{y}} = \mathbf{C}\hat{\mathbf{x}} \in \mathbb{R}^p$ is the output estimate, and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{L}$ are appropriately dimensioned real matrices. The first two terms of (3.10) are a copy of the process model given in (3.9). The last term of (3.10) is the innovation term and serves to drive $\hat{\mathbf{x}}$ toward $\mathbf{x}$. The selected innovation term is the canonical choice and the observer gain $\mathbf{L}$ is typically designed to place the eigenvalues of $\mathbf{A} - \mathbf{L}\mathbf{C}$ in the closed left half plane. The innovation, however, can also be thought

of as propagating the state estimate in the gradient descent direction of the function

$$f(\hat{\mathbf{x}}, \mathbf{x}) = \tfrac{1}{2}||\hat{\mathbf{x}} - \mathbf{x}||_2^2. \tag{3.11}$$

The gradient of f with respect to $\hat{\mathbf{x}}$ is $\boldsymbol{\nabla}_{\hat{\mathbf{x}}} f = \hat{\mathbf{x}} - \mathbf{x}$, and thus the Luenberger observer propagates the state estimate estimate in the gradient descent direction of the function f.

3.3.2 Cost Function and Gradient-Based Observer Design

Given measurements v_y and y, the goal is to design

$$\dot{\hat{X}} = F(\hat{X}, v_y, y), \tag{3.12}$$

such that $\hat{X} \to X$ as $t \to \infty$, where $\hat{X}$ is the estimate of X, and $F : G \times \mathfrak{g} \times \mathcal{M}^\ell \to TG$ is a parameterized vector field on G. The gradient-based design methodology considered in [14] operates on a similar principle to the Luenberger observer, in that it will be composed of a process model component and an innovation term based on the gradient descent direction of a cost function. Consequently, F has the form of

$$F(\hat{X}, v_y, y) = \hat{X} v_y + U(\hat{X}, y), \tag{3.13}$$

where $\hat{X} v_y \in T_{\hat{X}} G$ is a copy of the system dynamics (3.2) and $U : G \times \mathcal{M}^\ell \to TG$ is the innovation term.

As mentioned previously, the innovation will be chosen as the gradient descent direction of a cost function. To this end, let $g : \mathcal{M}^\ell \times \mathcal{M}^\ell \to \mathbb{R}^+$ be a smooth symmetric cost function such that $g(\hat{y}, y)$, where $\mathcal{I} = (I, \ldots, I) \in G^\ell$, describes the error between predicted observations $\hat{y} = h(\hat{X}, \mathcal{I}, \bar{y})$ and true observations $y = h(X, N, \bar{y})$. For the gradient-based observer, the gradient of g with respect to the state estimate $\hat{X}$ will be taken. Consequently, it will be useful to reparameterize the function g such that the function maps from the group G to $\mathbb{R}^+$, rather than from $\mathcal{M}^\ell$ to $\mathbb{R}^+$. This is not strictly necessary, see for example in [45] where an observer is constructed directly on the output space. However, in this dissertation the observers under consideration will evolve strictly on the group and therefore the reparameterization will be useful. The reparameterization can be accomplished in a

number of ways. First, consider the cost function $f_y : G \times \mathcal{M}^\ell \to \mathbb{R}^+$ defined as

$$\begin{aligned} f_y(\hat{X}, y) &= g(h(\hat{X}, \mathcal{I}, \bar{y}), h(X, N, \bar{y})) \\ &= g(h(\hat{X}, \mathcal{I}, \bar{y}), y). \end{aligned} \tag{3.14}$$

The second reparameterization considered is the function $f_N : G \times G \times G^\ell \to \mathbb{R}^+$ given by

$$f(\hat{X}, X, N) = g(h(\hat{X}, \mathcal{I}, \bar{y}), h(X, N, \bar{y})). \tag{3.15}$$

To simplify notation in later sections, it will be useful to define $f_N : G \times G \to \mathbb{R}^+$ as

$$f_N(\hat{X}, X) \triangleq f(\hat{X}, X, N). \tag{3.16}$$

When it is necessary to explicitly consider noise in the output measurement, as will be done in Chapters 6 and 7, the costs in (3.14) or (3.15) will be considered. Otherwise, most of the properties of the gradient-based observers considered in this dissertation will be derived with the assumption that $N = \mathcal{I}$. In this case, there is no multiplicative noise associated with the system outputs and (3.15) is expressed as

$$f_{\mathcal{I}}(\hat{X}, X) = g(h(\hat{X}, \mathcal{I}, \bar{y}), h(X, \mathcal{I}, \bar{y})). \tag{3.17}$$

The gradient-based observer can now be written as

$$\dot{\hat{X}} = \hat{X} v_y - k \nabla_{\hat{X}} f_y(\hat{X}, y), \tag{3.18}$$

where $k \in (0, \infty)$ is an observer gain [12, 14, 45]. Neglecting noise, (3.18) can be written as

$$\dot{\hat{X}} = \hat{X} v_y - k \nabla_{\hat{X}} f_{\mathcal{I}}(\hat{X}, X). \tag{3.19}$$

A pictorial representation of the gradient-based observer is shown in Fig. 3.1. It will often be useful to represent the gradient of the cost as an element of the Lie algebra. Recall that an element of the tangent space at any point $X \in G$ can be mapped to the Lie algebra algebra by either the mapping $T_X R_{X^{-1}}(\cdot)$ or $T_X L_{X^{-1}}(\cdot)$. Doing so the gradient $\nabla_{\hat{X}} f_y(\hat{X}, y) \in T_{\hat{X}} G$ can be represented by either

$$e_R = T_{\hat{X}} R_{\hat{X}^{-1}}(\nabla_{\hat{X}} f_y(\hat{X}, y)) \in \mathfrak{g}, \tag{3.20}$$

or

$$e_L = T_{\hat{X}} L_{\hat{X}^{-1}}(\nabla_{\hat{X}} f_y(\hat{X}, y)) \in \mathfrak{g}. \tag{3.21}$$

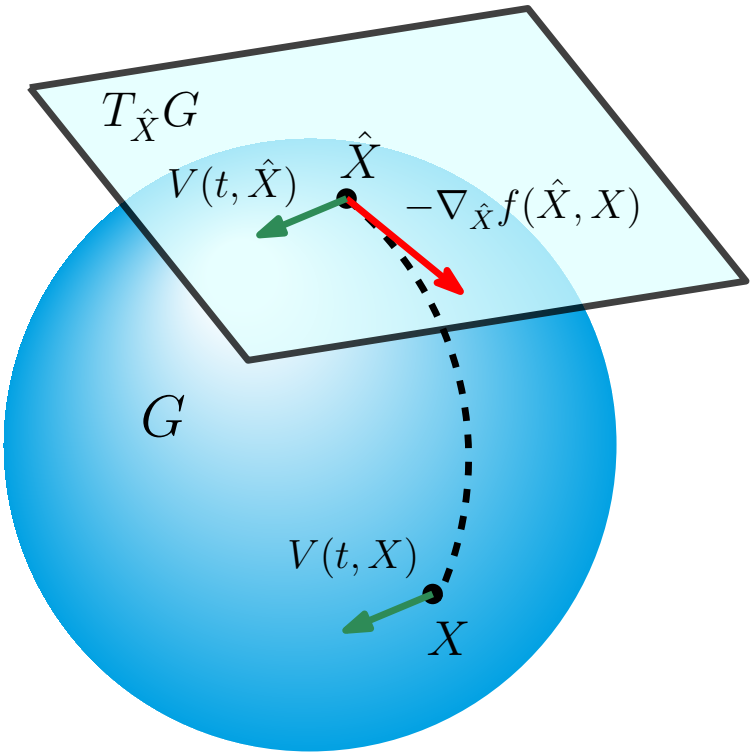


Figure 3.1: The gradient-based observer. The gradient-based observer is composed of two parts. The first copies the system kinematics, represented here as a green arrow, while the second is an innovation term, shown as a red arrow. The innovation term propagates the state estimate toward the true state along a curve, shown as a dashed line, dictated by the gradient-descent direction of a cost function.

In other words, the gradient can be written as

$$\nabla_{\hat{X}} f_y(\hat{X}, y) = e_R \hat{X} = \hat{X} e_L, \tag{3.22}$$

where the relationship between e_R and e_L is given by $e_R = \mathrm{Ad}_{\hat{X}}(e_L)$. Consequently, the gradient-based observer can be written equivalently as either

$$\dot{\hat{X}} = \hat{X} v_y - k e_R \hat{X}, \tag{3.23}$$

or

$$\dot{\hat{X}} = \hat{X} v_y - k \hat{X} e_L. \tag{3.24}$$

3.4 Error Dynamics

In order to assess the stability guarantees of the gradient-based observer in later chapters it will be necessary to determine the error dynamics associated with the observer. The group error may be defined by either right or left translation of $\hat{X}$ by X^{-1} giving $\tilde{X}_R \triangleq \hat{X}X^{-1}$ or $\tilde{X}_L \triangleq X^{-1}\hat{X}$ [14]. The following lemmas and definitions will be useful.

Lemma 3.1 (Lemma 10 [14])**.** *Let $X : \mathbb{R} \to G$ and $Y : \mathbb{R} \to G$ be two smooth curves on G. Then,*

$$\dot{(YX^{-1})} = T_Y R_{X^{-1}} \dot{Y} - T_{X^{-1}} L_Y T_I L_{X^{-1}} T_X R_{X^{-1}} \dot{X}, \tag{3.25}$$

and

$$\dot{(X^{-1}Y)} = -T_{X^{-1}} R_Y T_I L_{X^{-1}} T_X R_{X^{-1}} \dot{X} + T_Y L_{X^{-1}} \dot{Y}. \tag{3.26}$$

Definition 3.2 (Right- and Left-Invariant Cost Functions [14])**.** A cost function $f : G \times G \to \mathbb{R}^+$ is *right-invariant* if for all $X, Y, Z \in G$

$$f(X, Y) = f(XZ, YZ), \tag{3.27}$$

and is *left-invariant* if for all $X, Y, Z \in G$

$$f(X, Y) = f(ZX, ZY). \tag{3.28}$$

Definition 3.3 (Bi-Invariant Cost Function)**.** A cost function $f : G \times G \to \mathbb{R}^+$ is *bi-invariant* if it is both right-invariant and left-invariant.

Definition 3.4 (Right- and Left- Invariant Riemannian Metric [17])**.** A Riemannian metric $\langle \cdot, \cdot \rangle_X$ is right-invariant if

$$\langle V, U \rangle_X = \langle T_X R_Y(V), T_X R_Y(U) \rangle_{XY} \tag{3.29}$$

for all $X, Y \in G$ and $V, U \in T_X G$, and is left-invariant if

$$\langle V, U \rangle_X = \langle T_X L_Y(V), T_X L_Y(U) \rangle_{YX} \tag{3.30}$$

for all $X, Y \in G$ and $V, U \in T_X G$.

Definition 3.5 (Bi-Invariant Riemannian Metric)**.** A Riemannian metric is *bi-invariant* if it is both right-invariant and left-invariant.

The following lemma, proved in [14], is quite important in showing that the error dynamics associated with the gradient-based observer are autonomous, provided the cost and metric satisfy some invariance properties.

Lemma 3.6 (Lemma 16 [14]). *Let $f : G \times G \to \mathbb{R}$ be a right-invariant cost function and take a right-invariant Riemannian metric. Then, for all $X, Y, Z \in G$*

$$T_X R_{Z^{-1}} \nabla_X f(X, YZ) = \nabla_{XZ^{-1}} f(XZ^{-1}, Y). \tag{3.31}$$

If f is left-invariant and a left-invariant Riemannian metric is taken then,

$$T_X L_{Y^{-1}} \nabla_X f(X, YZ) = \nabla_{XZ^{-1}} f(Y^{-1}X, Z). \tag{3.32}$$

As a result of Lemma 3.6, the gradient of a right- (or left-) invariant cost function taken with respect to a right- (or left-) invariant metric can be written as a function of only the right (or left) group errors $\tilde{X}_R$ (or $\tilde{X}_L$). This can be seen by setting $X = \hat{X}$, $Y = I$, and $Z = X$ in (3.31) or $X = \hat{X}$, $Y = X$, and $Z = I$ in (3.32). It is now possible to determine the error dynamics associated with gradient-based observer.

Theorem 3.7 (Theorem 17 [14]). *Consider the system dynamics given in* (3.2)*. Assume that there is no noise associated with the group velocity measurement and partial state measurements such that $d = 0$ and $N = \mathcal{I}$. Let $f_{\mathcal{I}}$ be a right-invariant cost function and take a right-invariant Riemannian metric on G. Then, the error dynamics associated with the gradient-based observer* (3.19) *with group error $\tilde{X}_R = \hat{X}X^{-1}$ is given by*

$$\dot{\tilde{X}}_R = -k \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I). \tag{3.33}$$

A proof of Theorem 3.7 can be found in [14]. The results of Theorem 3.7 state that the error dynamics associated with the gradient-based observer are autonomous provided the system dynamics are left-invariant and the cost and Riemannian metric are right-invariant. A similar result is also shown in [14] for the case of right-invariant system dynamics coupled with a left-invariant cost and Riemannian metric. In this dissertation, only the case of left-invariant system dynamics are considered. This is due to the fact that it is assumed that the group velocity v is available for measurement rather than the equivalent velocity for right-invariant system dynamics, $w \in \mathfrak{g}$, $\dot{X} = wX$. For the Lie group $SO(3)$ this is equivalent to assuming that angular velocity in the body frame is available for measurement, rather than in the global or inertial frame.

Chapter 4

Extensions to Attitude Estimation

4.1 Introduction

Robotic vehicles require accurate knowledge of their orientation to execute autonomous maneuvers. However, orientation, also referred to as attitude, cannot be directly measured. Instead, the attitude of a vehicle must be ascertained from a variety of available measurements, many of which may be of poor quality. This is the problem of attitude estimation. The attitude estimation of a miniature aerial vehicle, for example, can be a difficult task. Such vehicles are typically equipped with low-cost inertial measurement units (IMUs). These low-cost sensors are characterized by poor noise properties that can complicate attitude estimation [10, 22, 46, 47].

Historically, attitude estimation has been accomplished using a form of extended Kalman filter (EKF) [48], a nonlinear variant of the popular Kalman filter [49]. The multiplicative EKF (MEKF), in particular, has been widely successful in the field of spacecraft attitude estimation [50]. However, it has been found that traditional Kalman filtering techniques can be difficult to apply robustly to vehicles with poor quality sensors [10, 22]. In addition, the non-Gaussian noise associated with these sensors [22] and the requisite linearizations inherent in the EKF formulation [21] can result in poor estimator performance [22]. Other estimation techniques, unscented filters [51–54] and particle filters [55–58] for example, can be difficult to apply on low-cost vehicles as well. These methods may be too computationally demanding to be implemented on vehicles with limited processing resources [47].

The poor performance of traditional estimation techniques when applied on low-cost vehicles has resulted in the popularity of a nonlinear attitude estimator [10]. The attitude estimator of [10] can be interpreted as a gradient-based observer on the Lie

group $SO(3)$. As was shown in Chapter 3, the design of the gradient-based observer is inherently linked to the selection of a cost function. Therefore, the behavior and properties of the estimator can often be derived directly from this function. One such popular cost function is considered in [22] and [10]. This same cost function is used in the stability analyses presented in [24, 25, 59], and as a consequence these estimators share similar convergence properties. A variation of the cost is used in the stability analysis of [60] and in the objective function of [61]. A drawback of the estimator first considered in [22] is that the attitude estimator requires an instantaneous algebraic construction of the attitude. This can be disadvantageous because it introduces additional floating point operations and can cause numerical difficulties [47]. This was rectified in [47] and [10], where a nonlinear attitude estimator based directly on measurements is considered. Again, the convergence properties of these estimators are related to the choice of a particular cost function, one with the same underlying form as those found in [22] and [60]. Similar estimators that use the same cost function, and therefore have similar convergence properties, can be found in [62] and [23].

The cost function of [22], and its variations [10, 47, 60], have seen widespread use in the context of rigid-body attitude control [17]. An alternate cost function has been considered in [63] for the purpose of trajectory tracking control. This alternative attitude error function yields a controller that remains aggressive for large attitude error, unlike those previously considered, resulting in superior convergence properties. Similarly, an alternate attitude cost function is considered in [64] in the context of spacecraft attitude control.

This chapter's main contribution is the introduction of two alternate cost functions, inspired by [63] and [64], that, when used in the observer design methodology of [14, 16] discussed in Chapter 3, yields attitude observers with superior convergence properties when compared to similar observers previously developed. The proposed observers belongs to the same class of nonlinear observers as [10, 22–25, 46, 47, 59–62, 65], however, the different cost functions used result in innovation terms that remains monotonically increasing with respect to attitude estimation error. In practical terms, the proposed observers allow for faster convergence for large initialization errors all the while maintaining equivalent steady-state performance when compared to previous attitude observers. The chapter begins with preliminaries in Sec. 4.2, and introduces the alternate cost functions in Sec. 4.3. Observer design on $SO(3)$ is discussed in Sec. 4.4 and a numerical example is presented in Sec. 4.5. Finally, closing remarks are given in Sec. 4.6.

4.2 Preliminaries

4.2.1 Notation

A frame of reference is defined as a set of three orthonormal dextral basis vectors [66]. Given a reference frame, a matrix composed of the basis vectors associated with the frame is formed, called a vectrix, defined as [66]

$$\underrightarrow{\mathcal{F}}_a = \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix}. \tag{4.1}$$

A physical vector, denoted $\underrightarrow{r}$, may be resolved in any frame. Given frames $\mathcal{F}_a$ and $\mathcal{F}_b$, $\underrightarrow{r}$ may be resolved as either $\underrightarrow{r} = \underrightarrow{\mathcal{F}}_a{}^{\mathsf{T}}\mathbf{r}_a = \underrightarrow{\mathcal{F}}_b{}^{\mathsf{T}}\mathbf{r}_b$, where $\mathbf{r}_a, \mathbf{r}_b \in \mathbb{R}^3$. The direction cosine matrix (DCM) that transforms a vector resolved in $\mathcal{F}_a$ to $\mathcal{F}_b$ is defined as $\mathbf{C}_{ba} \triangleq \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_a{}^{\mathsf{T}}$ such that $\mathbf{r}_b = \mathbf{C}_{ba}\mathbf{r}_a$. Alternatively, $\mathbf{r}_a = \mathbf{C}_{ab}\mathbf{r}_b$, where $\mathbf{C}_{ab} = \mathbf{C}_{ba}^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b{}^{\mathsf{T}}$. Direction cosine matrices constitute the Lie group of rigid body rotations denoted $SO(3)$, where $SO(3) = \{\mathbf{C} \in SO(3) \mid \mathbf{C}^{\mathsf{T}}\mathbf{C} = \mathbf{1},\ \det \mathbf{C} = +1\}$. In the robotics community the attitude of a rigid body is often identified by $\mathbf{R} \triangleq \mathbf{C}_{ab} \in SO(3)$ [67]. In this dissertation, notation will often switch between $\mathbf{R}$ and $\mathbf{C}_{ab}$ depending on whether it is necessary to be explicit on the frames of reference under consideration.

The Lie algebra of $SO(3)$ is denoted $\mathfrak{so}(3)$, where $\mathfrak{so}(3) = \{\mathbf{S} \in \mathbb{R}^{3\times 3} \mid \mathbf{S} = -\mathbf{S}^{\mathsf{T}}\}$. A Riemannian metric on $SO(3)$ is required. Consider first an inner product on $\mathfrak{so}(3)$, denoted $\langle \cdot, \cdot \rangle : \mathfrak{so}(3) \times \mathfrak{so}(3) \to \mathbb{R}$, defined as the standard matrix inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}^{\mathsf{T}}\mathbf{B})$ for all $\mathbf{A}, \mathbf{B} \in \mathfrak{so}(3)$. Then, a Riemannian metric on $SO(3)$ may be defined as $\langle \cdot, \cdot \rangle_{\mathbf{R}} : T_{\mathbf{R}}SO(3) \times T_{\mathbf{R}}SO(3) \to \mathbb{R}$ as [17]

$$\langle \mathbf{V}, \mathbf{U} \rangle_{\mathbf{R}} = \langle \mathbf{V}\mathbf{R}^{\mathsf{T}}, \mathbf{U}\mathbf{R}^{\mathsf{T}} \rangle,\ \forall \mathbf{V}, \mathbf{U} \in T_{\mathbf{R}}SO(3). \tag{4.2}$$

4.2.2 Attitude Kinematics

Consider a rigid body and two frames of reference, $\mathcal{F}_a$ and $\mathcal{F}_b$, where $\mathcal{F}_a$ is a datum frame and $\mathcal{F}_b$ is a body-fixed frame. Rigid-body kinematics are described by Poisson's kinematical equation, referred throughout simply as Poisson's equation,

$$\dot{\mathbf{C}}_{ba} = -\boldsymbol{\omega}_b^{ba\times}\mathbf{C}_{ba}, \tag{4.3}$$

where $\boldsymbol{\omega}_b^{ba} \in \mathbb{R}^3$ denotes the angular velocity of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ resolved in $\mathcal{F}_b$. The operator $(\cdot)^\times$ is a mapping $(\cdot)^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$, such that

$$\mathbf{r}^\times = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_3 & 0 \end{bmatrix}, \quad \forall \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \in \mathbb{R}^3.$$

The reverse operation is denoted $(\cdot)^\vee : \mathfrak{so}(3) \to \mathbb{R}^3$, such that $(\mathbf{r}^\times)^\vee = \mathbf{r}$, $\forall \mathbf{r} \in \mathbb{R}^3$. In terms of $\mathbf{R}$, (4.3) becomes

$$\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\omega}^\times, \tag{4.4}$$

where the superscripts and subscripts have been dropped for simplicity in notation.

The identities

$$\begin{aligned} -\mathbf{r}^\times \mathbf{s}^\times &= (\mathbf{r}^\mathsf{T}\mathbf{s})\mathbf{1} - \mathbf{s}\mathbf{r}^\mathsf{T}, & (4.5)\\ (\mathbf{r}^\times \mathbf{s})^\times &= \mathbf{s}\mathbf{r}^\mathsf{T} - \mathbf{r}\mathbf{s}^\mathsf{T}, & (4.6)\\ \mathbf{U}^\mathsf{T}\mathbf{r}^\times + \mathbf{r}^\times \mathbf{U} &= \big((\mathrm{tr}(\mathbf{U})\mathbf{1} - \mathbf{U})\mathbf{r}\big)^\times, & (4.7)\\ \tfrac{1}{2}\mathrm{tr}(\mathbf{U}\mathbf{r}^\times) &= -\mathbf{r}^\mathsf{T}\boldsymbol{\mathcal{P}}_a(\mathbf{U})^\vee, & (4.8)\\ \mathbf{r}^\mathsf{T}\mathbf{s} &= -\tfrac{1}{2}\mathrm{tr}(\mathbf{r}^\times \mathbf{s}^\times), & (4.9) \end{aligned}$$

where $\mathbf{r}, \mathbf{s} \in \mathbb{R}^3$, $\mathbf{U} \in \mathbb{R}^{3\times 3}$, and $\boldsymbol{\mathcal{P}}_a(\mathbf{U}) = \frac{1}{2}(\mathbf{U} - \mathbf{U}^\mathsf{T}) \in \mathbb{R}^{3\times 3}$, will be used in subsequent derivations [10, 63, 66]. The operator $\boldsymbol{\mathcal{P}}_a(\cdot)$ is the anti-symmetric projection operator. The symmetric projection operator is defined by $\boldsymbol{\mathcal{P}}_s(\mathbf{U}) = \frac{1}{2}(\mathbf{U} + \mathbf{U}^\mathsf{T})$. The following inequalities will also be useful, and hold for all $\mathbf{r} \in \mathbb{R}^3$, $\mathbf{A} = \mathbf{A}^\mathsf{T} \geq 0$, and $\mathbf{B} = \mathbf{B}^\mathsf{T} > 0$ [40, pp. 551-552]:

$$\begin{aligned} (\mathbf{r}^\mathsf{T}\mathbf{A}\mathbf{r})^2 &\leq \mathbf{r}^\mathsf{T}\mathbf{r}(\mathbf{r}^\mathsf{T}\mathbf{A}^2\mathbf{r}), & (4.10)\\ (\mathbf{r}^\mathsf{T}\mathbf{r})(\mathbf{r}^\mathsf{T}\mathbf{B}^2\mathbf{r}) &\leq (\mathbf{r}^\mathsf{T}\mathbf{B}\mathbf{r})^2 + \tfrac{1}{4}(\bar{\lambda}(\mathbf{B}) - \underline{\lambda}(\mathbf{B}))^2\mathbf{r}^\mathsf{T}\mathbf{r}, & (4.11) \end{aligned}$$

where $\underline{\lambda}(\cdot)$ and $\bar{\lambda}(\cdot)$ respectively denote the minimum and maximum eigenvalues of a matrix.

4.2.3 Estimation Error and Measurements

The goal is to estimate the attitude $\mathbf{R}$. For this purpose, let $\hat{\mathbf{R}} \in SO(3)$ be the estimate of $\mathbf{R}$ and define the group error as $\tilde{\mathbf{R}} = \hat{\mathbf{R}}\mathbf{R}^\mathsf{T}$ [10]. The goal of attitude estimation is to accurately determine the orientation of a body from available measurements. In this chapter, the attitude estimation of a rigid body equipped with

sensors that provide an angular velocity measurement along with two or more unit-length vector measurements is considered. For example, a typical IMU is equipped with an accelerometer and magnetometer. The magnetometer measures Earth's local magnetic field vector and, assuming the acceleration of the body remains small, the accelerometer may be used to measure Earth's local gravitational field vector [10].

The rate-gyro measures

$$\boldsymbol{\omega}_y = \boldsymbol{\omega} + \mathbf{b} + \boldsymbol{\mu}, \tag{4.12}$$

where $\boldsymbol{\omega}_b^y$ is the measured angular velocity, $\mathbf{b}$ is the rate-gyro bias, and $\boldsymbol{\mu}$ is the measurement noise associated with the rate gyroscope. Other sources of error (e.g., sensor alignment errors) are not considered.

It is also assumed that there are n vector measurements available where $n \in \mathbb{Z}$, $n \geq 2$. Each vector measurement measures a unit-length reference vector whose components resolved in $\mathcal{F}_a$ are constant and known. For example, over short distances and time intervals Earth's magnetic field vector, resolved in $\mathcal{F}_a$, may be considered constant and can be computed using a Geomagnetic field model [68]. The unit-length vector measurements are measured in the body frame as $\mathbf{y}_b^j = \mathbf{R}^\mathsf{T}\mathbf{y}_a^j$, $j = 1, \ldots, n$, where $\mathbf{y}_a^j \in \mathbb{S}^2$, $\mathbb{S}^2 = \{\mathbf{r} \in \mathbb{R}^3 \mid \sqrt{\mathbf{r}^\mathsf{T}\mathbf{r}} = 1\}$, is the j^{th} reference vector resolved in the inertial frame. Let $\mathbf{y}_e^j = \hat{\mathbf{R}}^\mathsf{T}\mathbf{y}_a^j \in \mathbb{S}^2$ be the j^{th} estimated measurement resolved in the body frame. The error between the j^{th} estimated measurement and the j^{th} measurement is defined as [10, 47]

$$\begin{aligned} E_j(\hat{\mathbf{R}}, \mathbf{R}) &= (\mathbf{y}_e^j - \mathbf{y}_b^j)^\mathsf{T}(\mathbf{y}_e^j - \mathbf{y}_b^j) \\ &= 1 - \mathbf{y}_e^{j^\mathsf{T}}\mathbf{y}_b^j \\ &= 1 - \mathrm{tr}(\mathbf{R}^\mathsf{T}\mathbf{y}_a^j\mathbf{y}_a^{j^\mathsf{T}}\hat{\mathbf{R}}) \\ &= 1 - \mathrm{tr}(\mathbf{y}_a^j\mathbf{y}_a^{j^\mathsf{T}}\hat{\mathbf{R}}\mathbf{R}^\mathsf{T}). \end{aligned}$$

The total estimation error associated with all measurements is then defined as

$$\begin{aligned} E(\hat{\mathbf{R}}, \mathbf{R}) &= \tfrac{1}{2}\sum_{j=1}^{n} k_j E_j(\hat{\mathbf{R}}, \mathbf{R}) \\ &= \tfrac{1}{2}\left(\sum_{j=1}^{n} k_j - \mathrm{tr}(\mathbf{M}_a\hat{\mathbf{R}}\mathbf{R}^\mathsf{T})\right), \end{aligned} \tag{4.13}$$

where $\mathbf{M}_a = \mathbf{M}_a^\mathsf{T} = \sum_{j=1}^{n} k_j\mathbf{y}_a^j\mathbf{y}_a^{j^\mathsf{T}}$, and $0 < k_j < \infty$, $j = 1, \ldots, n$, are positive constants. The trace of $\mathbf{M}_a$ is $\mathrm{tr}(\mathbf{M}_a) = \sum_{j=1}^{n} k_j$, and as a result write $E(\hat{\mathbf{R}}, \mathbf{R}) =$

$\text{tr}\big(\mathbf{M}_a(\mathbf{1} - \hat{\mathbf{R}}\mathbf{R}^\mathsf{T})\big)$. For simplicity, E may be expressed as

$$E(\hat{\mathbf{R}}, \mathbf{R}) = \tilde{E}(\tilde{\mathbf{R}}) = \tfrac{1}{2}\text{tr}(\mathbf{M}_a(\mathbf{1} - \tilde{\mathbf{R}})), \tag{4.14}$$

where $\tilde{E} : SO(3) \to \mathbb{R}^+$.

It will also be useful to consider the matrix $\mathbf{G}_a = \text{tr}(\mathbf{M}_a)\mathbf{1} - \mathbf{M}_a \in \mathbb{R}^{3\times 3}$. The following statements regarding $\mathbf{G}_a$ hold, provided $\text{rank}(\mathbf{M}_a) \geq 2$, [17, p. 553]

(i) $\mathbf{G}_a$ is symmetric and positive definite.

(ii) The eigenvectors of $\mathbf{M}_a$ coincide with the eigenvectors of $\mathbf{G}_a$.

(iii) Let $\lambda(\cdot)$ denote the spectrum of a matrix. If $\lambda(\mathbf{M}_a) = \{\lambda_1, \lambda_2, \lambda_3\}$, then $\lambda(\mathbf{G}_a) = \{\lambda_2 + \lambda_3, \lambda_1 + \lambda_3, \lambda_1 + \lambda_2\}$.

For the remainder of this chapter, it is assumed that $\text{rank}(\mathbf{M}_a) \geq 2$ and thus the above statements hold. This indicates that at least two linearly independent reference vectors are available for measurement. Also, let $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3\}$ be an orthonormal set of eigenvectors of $\mathbf{M}_a$ and $\mathbf{G}_a$ associated with eigenvalues $\{\lambda_1,\ \lambda_2,\ \lambda_3\}$ and $\{\lambda_2 + \lambda_3, \lambda_1 + \lambda_3, \lambda_1 + \lambda_2\}$, respectively, where it is assumed that $\lambda_3 \leq \lambda_2 \leq \lambda_1$ [17, p. 553].

4.3 Cost Function Selection

To construct the gradient-based observer for $SO(3)$, a cost function will first need to be selected. Since the cost function is directly related to the innovation term, its selection is crucial in determining the behavior of the observer. Given a cost function on $SO(3)$, $f : SO(3) \times SO(3) \to \mathbb{R}$, its gradient is expressed as

$$\nabla_{\hat{\mathbf{R}}} f(\hat{\mathbf{R}}, \mathbf{R}) = \mathbf{e}^\times \hat{\mathbf{R}}, \tag{4.15}$$

where $\mathbf{e} \in \mathbb{R}$ is referred to as the attitude error vector, or innovation, associated with f. In this section, the properties of three cost functions on $SO(3)$ and their corresponding error vectors are explored. The first cost function that is considered is simply the measurement error E described in (4.14). This is a well known cost function that is often used in the context of rigid-body attitude control and estimation problems. In fact, the use of E as the cost function in the gradient-based observer on $SO(3)$ results in the popular attitude estimator first proposed in [10], referred to as a nonlinear complementary filter. Unfortunately, the use of E as a cost function has a severe drawback in that it results in slow convergence of the attitude estimate,

especially in cases when the attitude estimate is poorly initialized. In Sec. 4.3.1 the properties of E as a cost function are explored. Then, two alternate cost functions are explored in Secs. 4.3.2 and 4.3.3. A comparison of the three cost functions is given in Sec. 4.3.4, where the slow convergence properties associated with E is explained.

4.3.1 Traditional Cost Function

Consider the cost function $\gamma : SO(3) \times SO(3) \to \mathbb{R}^+$ given by

$$\gamma(\hat{\mathbf{R}}, \mathbf{R}) = \mathrm{tr}(\mathbf{M}_a(\mathbf{1} - \hat{\mathbf{R}}\mathbf{R}^\mathsf{T})), \tag{4.16}$$

or alternatively

$$\tilde{\gamma}(\tilde{\mathbf{R}}) = \mathrm{tr}(\mathbf{M}_a(\mathbf{1} - \tilde{\mathbf{R}})), \tag{4.17}$$

where $\gamma(\hat{\mathbf{R}}, \mathbf{R}) = \tilde{\gamma}(\tilde{\mathbf{R}})$. The cost function γ is simply the total measurement error E given in (4.14) scaled by a factor of 2. The factor of 2 is added to simplify the form of the resulting error vector and is a matter of convenience. The properties of γ are well known and understood, see for example in [17] or [10]. However, it will be useful for the purpose of comparison to alternate cost functions to explore these properties. The properties of γ are described in the following proposition.

Proposition 4.1. *Consider the attitude error function $\tilde{\gamma} : SO(3) \to \mathbb{R}^+$ given by*

$$\tilde{\gamma}(\tilde{\mathbf{R}}) = \mathrm{tr}(\mathbf{M}_a(\mathbf{1} - \tilde{\mathbf{R}})), \tag{4.18}$$

and associated attitude error vector

$$\mathbf{e}_\gamma = \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\mathsf{V}. \tag{4.19}$$

The attitude error function $\tilde{E}$ has the following properties.

(i) $\tilde{\gamma}$ is positive definite about $\tilde{\mathbf{R}} = \mathbf{1}$.

(ii) The gradient of $\tilde{\gamma}$ with respect to the Riemannian metric defined in (4.2) is $\nabla\tilde{\gamma}(\tilde{\mathbf{R}}) = \mathbf{e}_\gamma^\times \tilde{\mathbf{R}}$.

(iii) Let $\mathscr{L} = \{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\gamma}(\tilde{\mathbf{R}}) < 2\underline{\lambda}(\mathbf{G}_a)\}$. The only critical point of $\tilde{\gamma}$ in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$.

(iv) For all $\tilde{\mathbf{R}} \in SO(3)$ there exists constants b_1 and b_2 such that

$$b_1||\mathbf{e}_\gamma||_2^2 \leq \gamma(\tilde{\mathbf{R}}) \leq b_2||\mathbf{e}_\gamma||_2^2, \tag{4.20}$$

where

$$b_1 = \frac{8\underline{\lambda}(\mathbf{G}_a)}{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2} \tag{4.21}$$

and

$$b_2 = \frac{2}{\underline{\lambda}(\mathbf{G}_a)}. \tag{4.22}$$

Proof. Properties (i) through (iii) are proved in [17, p.553]. However, these results are included here for completeness. First, express $\tilde{\mathbf{R}}$ as

$$\tilde{\mathbf{R}} = \exp(\phi_e \mathbf{a}_e^{\times}) = \cos(\phi_e)\mathbf{1} + (1 - \cos(\phi_e))\mathbf{a}_e\mathbf{a}_e^{\mathsf{T}} + \sin(\phi_e)\mathbf{a}_e^{\times}, \tag{4.23}$$

where $\mathbf{a}_e \in \mathbb{S}^2$ and $\phi_e \in \mathbb{R}$ are the axis and angle of Euler's theorem [66, 67]. Then, (4.14) may be written as

$$\begin{aligned}
\tilde{E} &= \tfrac{1}{2}\mathrm{tr}\big((1 - \cos(\phi))\mathbf{M}_a(\mathbf{a}_e^{\mathsf{T}}\mathbf{a}_e\mathbf{1} - \mathbf{a}_e\mathbf{a}_e^{\mathsf{T}}) - \sin(\phi_e)\mathbf{M}_a\mathbf{a}_e^{\times}\big) \\
&= \tfrac{1}{2}\mathrm{tr}\big(- (1 - \cos(\phi))\mathbf{M}_a\mathbf{a}_e^{\times}\mathbf{a}_e^{\times} - \sin(\phi_e)\mathbf{M}_a\mathbf{a}_e^{\times}\big) \\
&= \tfrac{1}{2}\mathrm{tr}\big([-(1 - \cos(\phi))\mathbf{M}_a\mathbf{a}_e^{\times} - \sin(\phi_e)\mathbf{M}_a]\mathbf{a}_e^{\times}\big) \\
&= -\mathbf{a}_e^{\mathsf{T}}\boldsymbol{\mathcal{P}}_a(-(1 - \cos(\phi))\mathbf{M}_a\mathbf{a}_e^{\times} - \sin(\phi_e)\mathbf{M}_a)^{\mathsf{V}} \\
&= \tfrac{1}{2}(1 - \cos(\phi_e))\mathbf{a}_e^{\mathsf{T}}(\mathbf{M}_a\mathbf{a}_e^{\times} + \mathbf{a}_e^{\times}\mathbf{M}_a)^{\mathsf{V}} \\
&= \tfrac{1}{2}(1 - \cos(\phi_e))\mathbf{a}_e^{\mathsf{T}}(\mathrm{tr}(\mathbf{M}_a)\mathbf{1} - \mathbf{M}_a)\mathbf{a}_e \\
&= \sin^2(\phi_e/2)\mathbf{a}_e^{\mathsf{T}}\mathbf{G}_a\mathbf{a}_e.
\end{aligned} \tag{4.24}$$

Thus, it follows that $\tilde{\gamma}(\tilde{\mathbf{R}}) = 2\sin^2(\phi_e/2)\mathbf{a}_e^{\mathsf{T}}\mathbf{G}_a\mathbf{a}_e$, and $\tilde{\gamma}(\tilde{\mathbf{R}}) \geq 0$ and $\tilde{\gamma}(\tilde{\mathbf{R}}) = 0$ if and only if $\phi_e \in \{0,\ \pm 2\pi,\ \pm 4\pi, \ldots\}$, which implies that $\tilde{\mathbf{R}} = \mathbf{1}$. This proves (i).

Item (ii) can be proved by directly computing the gradient of $\tilde{\gamma}(\tilde{\mathbf{R}})$. Let $\boldsymbol{\Gamma}(\epsilon)$ be a curve at $\tilde{\mathbf{R}}$ given by $\boldsymbol{\Gamma}(\epsilon) = \exp(\epsilon\mathbf{S})\tilde{\mathbf{R}}$, where $\epsilon \in \mathbb{R}$ and $\mathbf{S} \in \mathfrak{so}(3)$. Then, the gradient of $\tilde{E}$ with respect to the Riemannian metric defined in (4.2) is defined as

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\gamma}(\boldsymbol{\Gamma}(\epsilon))\bigg|_{\epsilon=0} = \langle \nabla\tilde{\gamma}(\tilde{\mathbf{R}}), \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}. \tag{4.25}$$

From the definition of $\tilde{\gamma}$ and $\mathbf{\Gamma}$ it follows that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\gamma}(\mathbf{\Gamma}(\epsilon))\bigg|_{\epsilon=0} &= -\mathrm{tr}(\tilde{\mathbf{R}}\mathbf{M}_a\mathbf{S}) && (4.26)\\
&= -\tfrac{1}{2}\mathrm{tr}(\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\mathbf{S}) && (4.27)\\
&= \mathrm{tr}(\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^{\mathsf{T}}\mathbf{S}) && (4.28)\\
&= \langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a), \mathbf{S}\rangle && (4.29)\\
&= \langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}}, \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}, && (4.30)
\end{aligned}$$

which implies that

$$\begin{aligned}
\nabla\tilde{\gamma}(\tilde{\mathbf{R}}) &= \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}} && (4.31)\\
&= \mathbf{e}_\gamma^{\times}\tilde{\mathbf{R}}. && (4.32)
\end{aligned}$$

This proves (ii).

To show that the only critical point in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$, observe that all critical points must satisfy $\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a) = \mathbf{0}$. Alternatively,

$$\tilde{\mathbf{R}}\mathbf{M}_a - \tilde{\mathbf{R}}^{\mathsf{T}}\mathbf{M}_a = \mathbf{0}.$$

The values of $\tilde{\mathbf{R}}$ that satisfy $\tilde{\mathbf{R}}\mathbf{M}_a - \tilde{\mathbf{R}}^{\mathsf{T}}\mathbf{M}_a = \mathbf{0}$ are $\tilde{\mathbf{R}} = \mathbf{1}$ and $\tilde{\mathbf{R}} = \exp(\pi\mathbf{v}^{\times})$, where [17, p. 553]

(a) $\mathbf{v} \in \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for distinct λ_1, λ_2, and λ_3,

(b) $\mathbf{v} \in \{\alpha\mathbf{v}_i + \beta\mathbf{v}_j \mid \alpha^2 + \beta^2 = 1\} \cup \mathbf{v}_k$ for $i, j, k \in \{1, 2, 3\}$, $\lambda_i = \lambda_j \neq \lambda_k$, and

(c) $\mathbf{v} \in \mathbb{S}^2$ for $\lambda_1 = \lambda_2 = \lambda_3$.

Now, evaluating (4.24) at the critical points it is observed that $\tilde{E}(\mathbf{1}) = 0$, and

(a) $\tilde{E}(\exp(\pi\mathbf{v}_1^{\times})) = \mathbf{v}_1^{\mathsf{T}}\mathbf{G}_a\mathbf{v}_1 = \underline{\lambda}(\mathbf{G}_a)$,

(b) $\tilde{E}(\exp(\pi\mathbf{v}^{\times})) = \underline{\lambda}(\mathbf{G}_a)$, where $\mathbf{v} \in \mathbb{S}^2$, and $\lambda_1 = \lambda_2 = \lambda_3$.

Turning to (4.40), it is seen that the only critical point of ψ that lies in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$.

For item (iv), an expression for $\mathbf{e}_\gamma^{\mathsf{T}}\mathbf{e}_\gamma$ is first sought. Applying identities (4.5), (4.6), and (4.7), $\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^{\mathsf{V}}$ can be expressed as

$$\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^{\mathsf{V}} = \tfrac{1}{2}\left((1 - \cos(\phi_e))\mathbf{a}_e^{\times}\mathbf{G}_a\mathbf{a}_e + \sin(\phi_e)\mathbf{G}_a\mathbf{a}_e\right). \qquad (4.33)$$

Next, substituting (4.33) into (4.19) yields

$$\mathbf{e}_\gamma \;=\; \tfrac{1}{2}(1-\cos(\phi_e))\mathbf{a}_e^\times \mathbf{G}_a\mathbf{a}_e + \tfrac{1}{2}\sin(\phi_e)\mathbf{G}_a\mathbf{a}_e \tag{4.34}$$

and therefore,

$$\begin{aligned}
\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma &= \tfrac{1}{4}\left(-(1-\cos(\phi_e))\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e^\times\mathbf{a}_e^\times\mathbf{G}_a\mathbf{a}_e + \sin^2(\phi_e)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e\right)\\
&= -\sin^4(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e^\times\mathbf{a}_e^\times\mathbf{G}_a\mathbf{a}_e + \sin^2(\phi_e/2)\cos^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e\\
&= -\sin^4(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a(-\mathbf{1}+\mathbf{a}_e\mathbf{a}_e^\mathsf{T})\mathbf{G}_a\mathbf{a}_e + \sin^2(\phi_e/2)\cos^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e\\
&= -\sin^4(\phi_e/2)(-\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e + (\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2) + \sin^2(\phi_e/2)\cos^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e\\
&= \sin^4(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e - \sin^4(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2 + \sin^2(\phi_e/2)\cos^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e\\
&= \sin^2(\phi_e/2)(\sin^2(\phi_e/2)+\cos^2(\phi_e/2))\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e - \sin^4(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2\\
&= \sin^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a^2\mathbf{a}_e - \sin^4(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2.
\end{aligned} \tag{4.35}$$

where identity (4.6) has been applied. Let us now find the constant b_1. From (4.11), (4.24), (4.35), and (4.55), it follows that

$$\begin{aligned}
\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma &\le \sin^2(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2 - \sin^4(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2\\
&\quad + \tfrac{1}{4}(\bar\lambda(\mathbf{G}_a)-\underline\lambda(\mathbf{G}_a))^2\sin^2(\phi_e/2)\\
&= \sin^2(\phi_e/2)\cos^2(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2 + \tfrac{1}{4}(\bar\lambda(\mathbf{G}_a)-\underline\lambda(\mathbf{G}_a))^2\sin^2(\phi_e/2)\\
&= \cos^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e\tilde E + \frac{(\bar\lambda(\mathbf{G}_a)-\underline\lambda(\mathbf{G}_a))^2\tilde E}{4\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e}\\
&\le \bar\lambda(\mathbf{G}_a)\tilde E + \frac{(\bar\lambda(\mathbf{G}_a)-\underline\lambda(\mathbf{G}_a))^2}{4\underline\lambda(\mathbf{G}_a)}\tilde E\\
&= \left(\bar\lambda(\mathbf{G}_a) + \frac{(\bar\lambda(\mathbf{G}_a)-\underline\lambda(\mathbf{G}_a))^2}{4\underline\lambda(\mathbf{G}_a)}\right)\tilde E\\
&= \frac{(\bar\lambda(\mathbf{G}_a)+\underline\lambda(\mathbf{G}_a))^2}{4\underline\lambda(\mathbf{G}_a)}\tilde E.
\end{aligned} \tag{4.36}$$

Therefore it follows that

$$b_1\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma \le b_1\frac{(\bar\lambda(\mathbf{G}_a)+\underline\lambda(\mathbf{G}_a))^2}{8\underline\lambda(\mathbf{G}_a)}\tilde\gamma.$$

Setting

$$b_1 = \frac{8\underline\lambda(\mathbf{G}_a)}{(\bar\lambda(\mathbf{G}_a)+\underline\lambda(\mathbf{G}_a))^2},$$

gives $b_1\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma \le \tilde\gamma$. It remains to find b_2 such that $b_2\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma \ge \tilde\gamma$. Returning to (4.35) and

applying (4.10), gives

$$\begin{aligned}
\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma &\geq \sin^2(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_b\mathbf{a}_e)^2 - \sin^4(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_b\mathbf{a}_e)^2 \\
&= \sin^2(\phi_e/2)\cos^2(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_b\mathbf{a}_e)^2 \\
&\geq \cos^2(\phi_e/2)\underline{\lambda}(\mathbf{G}_a)\tilde{E} \\
&\geq \underline{\lambda}(\mathbf{G}_a)\tilde{E}. \qquad (4.37)
\end{aligned}$$

It follows that

$$b_2\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma \geq \tfrac{1}{2}b_2\underline{\lambda}(\mathbf{G}_a)\tilde{\gamma},$$

and setting

$$b_2 = \frac{2}{\underline{\lambda}(\mathbf{G}_a)},$$

gives $b_2\mathbf{e}_\gamma^\mathsf{T}\mathbf{e}_\gamma \geq \tilde{\gamma}$. □

Unfortunately, the error vector associated with γ does not remain monotonically increasing with respect to attitude error. This can be seen from (4.34), where, when $\mathbf{G}_a = \mathbf{1}$, the error vector $\mathbf{e}_\gamma$ is given by $\mathbf{e}_\gamma = \frac{1}{2}\sin(\phi_e)\mathbf{a}$ and thus as ϕ_e approaches π, the error vector approaches $\mathbf{0}$. This is highly undesirable as the error vector is directly related to the innovation term in the gradient-based observer. Consequently, the use of $\mathbf{e}_\gamma$ will result in slow convergence of the attitude estimate when large estimation errors are present. In Sec. 4.3.2 and Sec. 4.3.3, two alternate cost functions are explored that result in much more aggressive error vectors.

4.3.2 Quaternion Cost Function

As was just seen the use of the cost function γ defined (4.16) as the cost function of the gradient-based observer will result in slow convergence of the attitude estimate. In this section, an alternate attitude error function is introduced that results in a gradient-based observer with faster convergence properties. The cost function is denoted $\psi : SO(3) \times SO(3) \to \mathbb{R}^+$, and is given by

$$\psi(\hat{\mathbf{R}}, \mathbf{R}) = 2\sqrt{\underline{\lambda}(\mathbf{G}_a)} - 2\sqrt{\underline{\lambda}(\mathbf{G}_a) - E(\hat{\mathbf{R}}, \mathbf{R})}. \qquad (4.38)$$

In terms of the group error $\tilde{\mathbf{R}}$, write $\psi(\hat{\mathbf{R}}, \mathbf{R}) = \tilde{\psi}(\tilde{\mathbf{R}})$, $\tilde{\psi} : SO(3) \to \mathbb{R}^+$, where

$$\tilde{\psi}(\tilde{\mathbf{R}}) = 2\sqrt{\underline{\lambda}(\mathbf{G}_a)} - 2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}. \qquad (4.39)$$

The proposed error function is denoted the quaternion attitude error function due to the fact that, when $\mathbf{G}_a = \mathbf{1}$, the resulting error vector is the vector part of the quaternion associated with $\tilde{\mathbf{R}}$. Useful properties of ψ are given in the following proposition.

Proposition 4.2. *Consider the attitude error function $\tilde{\psi} : D \to \mathbb{R}^+$ given by*

$$\tilde{\psi}(\tilde{\mathbf{R}}) = 2\sqrt{\underline{\lambda}(\mathbf{G}_a)} - 2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}, \tag{4.40}$$

and associated attitude error vector

$$\mathbf{e}_\psi = \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee, \tag{4.41}$$

where $D = \left\{ \tilde{\mathbf{R}} \in SO(3) \mid \tilde{E}(\tilde{\mathbf{R}}) \leq \underline{\lambda}(\mathbf{G}_a) \right\}$ is the allowable domain of $\tilde{\psi}$. The attitude error vector, $\mathbf{e}_\psi$, is well defined in the sublevel set

$$\mathscr{L} = \left\{ \tilde{\mathbf{R}} \in SO(3) \mid \tilde{\psi}(\tilde{\mathbf{R}}) < \sqrt{\underline{\lambda}(\mathbf{G}_a)} \right\}, \tag{4.42}$$

where $\mathscr{L} \subset D$ and $D \setminus \mathscr{L} = \{\tilde{\mathbf{R}} \in SO(3) \mid E(\tilde{\mathbf{R}}) = \underline{\lambda}(\mathbf{G}_a)\}$. The attitude error function $\tilde{\psi}$ has the following properties.

(i) $\tilde{\psi}$ is positive definite about $\tilde{\mathbf{R}} = \mathbf{1}$.

(ii) The gradient of $\tilde{\psi}$ with respect to the Riemannian metric defined in (4.2) is $\nabla\tilde{\psi}(\tilde{\mathbf{R}}) = \mathbf{e}_\psi^\times \tilde{\mathbf{R}}$.

(iii) The only critical point of $\tilde{\psi}$ in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$.

(iv) Let $\Omega_c = \{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\psi}(\tilde{\mathbf{R}}) \leq c\}$, where $c < 2\sqrt{\underline{\lambda}(\mathbf{G}_a)}$. Then, for all $\tilde{\mathbf{R}} \in \Omega_c$ there exists constants b_1 and b_2 such that

$$b_1 ||\mathbf{e}_\psi||_2^2 \leq \tilde{\psi}(\tilde{\mathbf{R}}) \leq b_2 ||\mathbf{e}_\psi||_2^2, \tag{4.43}$$

where

$$b_1 = \frac{4\sqrt{\underline{\lambda}(\mathbf{G}_a)}(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2}{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2} \tag{4.44}$$

and

$$b_2 = \frac{2}{\sqrt{\underline{\lambda}(\mathbf{G}_a)}}. \tag{4.45}$$

Proof Substituting (4.24) into (4.40) yields

$$\tilde{\psi}(\tilde{\mathbf{R}}) = 2\sqrt{\underline{\lambda}(\mathbf{G}_a)} - 2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \sin^2(\phi_e/2)\mathbf{a}_e^{\mathsf{T}}\mathbf{G}_a\mathbf{a}_e}. \tag{4.46}$$

Thus, it follows that $\tilde{\psi} \geq 0$ and $\tilde{\psi} = 0$ if and only if $\phi_e \in \{0,\ \pm 2\pi,\ \pm 4\pi, \ldots\}$, which implies that $\tilde{\mathbf{R}} = \mathbf{1}$. This proves (i).

To prove item (ii), let $\mathbf{\Gamma}(\epsilon)$ be a curve at $\tilde{\mathbf{R}}$ given by $\mathbf{\Gamma}(\epsilon) = \exp(\epsilon\mathbf{S})\tilde{\mathbf{R}}$, where $\epsilon \in \mathbb{R}$ and $\mathbf{S} \in \mathfrak{so}(3)$. Then, the gradient of $\tilde{\psi}$ with respect to the Riemannian metric defined in (4.2) is defined as

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\psi}(\mathbf{\Gamma}(\epsilon))\bigg|_{\epsilon=0} = \langle \nabla\tilde{\psi}(\tilde{\mathbf{R}}), \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}. \tag{4.47}$$

From the definition of $\tilde{\psi}$ and $\mathbf{\Gamma}$ it follows that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\psi}(\mathbf{\Gamma}(\epsilon))\bigg|_{\epsilon=0} &= \frac{-1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}}\mathrm{tr}(\mathbf{M}_a\mathbf{S}\tilde{\mathbf{R}}) && (4.48)\\
&= \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}}\langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a), \mathbf{S}\rangle && (4.49)\\
&= \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}}\langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}}, \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}, && (4.50)
\end{aligned}$$

which implies that

$$\begin{aligned}
\nabla\tilde{\psi}(\tilde{\mathbf{R}}) &= \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}}\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}} && (4.51)\\
&= \mathbf{e}_\psi^{\times}\tilde{\mathbf{R}}, && (4.52)
\end{aligned}$$

proving (ii).

To show that the only critical point in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$, observe that all critical points must satisfy $\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a) = \mathbf{0}$. Alternatively,

$$\tilde{\mathbf{R}}\mathbf{M}_a - \tilde{\mathbf{R}}^{\mathsf{T}}\mathbf{M}_a = \mathbf{0}.$$

The values of $\tilde{\mathbf{R}}$ that satisfy $\tilde{\mathbf{R}}\mathbf{M}_a - \tilde{\mathbf{R}}^{\mathsf{T}}\mathbf{M}_a = \mathbf{0}$ are $\tilde{\mathbf{R}} = \mathbf{1}$ and $\tilde{\mathbf{R}} = \exp(\pi\mathbf{v}^{\times})$, where [17, p. 553]

(a) $\mathbf{v} \in \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for distinct λ_1, λ_2, and λ_3,

(b) $\mathbf{v} \in \{\alpha \mathbf{v}_i + \beta \mathbf{v}_j \mid \alpha^2 + \beta^2 = 1\} \cup \mathbf{v}_k$ for $i, j, k \in \{1, 2, 3\}$, $\lambda_i = \lambda_j \neq \lambda_k$, and

(c) $\mathbf{v} \in \mathbb{S}^2$ for $\lambda_1 = \lambda_2 = \lambda_3$.

Evaluating (4.24) at the critical points it is observed that $\tilde{E}(\mathbf{1}) = 0$, and

(a) $\tilde{E}(\exp(\pi \mathbf{v}_1^{\times})) = \mathbf{v}_1^{\mathsf{T}} \mathbf{G}_a \mathbf{v}_1 = \underline{\lambda}(\mathbf{G}_a)$,

(b) $\tilde{E}(\exp(\pi \mathbf{v}^{\times})) = \underline{\lambda}(\mathbf{G}_a)$, where $\mathbf{v} \in \mathbb{S}^2$, and $\lambda_1 = \lambda_2 = \lambda_3$.

Turning to (4.40), the only critical point of ψ that lies in $\mathscr{L}$ is $\tilde{\mathbf{R}} = \mathbf{1}$.

For item (iv), note that

$$\mathbf{e}_\psi = \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}}} \mathbf{e}_\gamma, \tag{4.53}$$

and therefore from (4.36)

$$\mathbf{e}_\psi^{\mathsf{T}} \mathbf{e}_\psi \leq \frac{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2}{16 \underline{\lambda}(\mathbf{G}_a)(\underline{\lambda}(\mathbf{G}_a) - \tilde{E})} \tilde{E}. \tag{4.54}$$

By assumption $\tilde{\mathbf{R}} \in \Omega_c$ and therefore from the definition of Ω_c and $\tilde{\psi}$ it follows that

$$\tfrac{1}{4}(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2 \leq \underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}) \tag{4.55}$$

for all $\tilde{\mathbf{R}} \in \Omega_c$. This implies that

$$\mathbf{e}_\psi^{\mathsf{T}} \mathbf{e}_\psi \leq \frac{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2}{4 \underline{\lambda}(\mathbf{G}_a)(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2} \tilde{E} \tag{4.56}$$

for all $\tilde{\mathbf{R}} \in \Omega_c$. Let us now define $\tilde{\delta} = 2\sqrt{\underline{\lambda}(\mathbf{G}_a)} + 2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}}$ such that $\tilde{E} = \frac{1}{4}\tilde{\psi}\tilde{\delta}$ The maximum value of $\tilde{\delta}$ in $\mathscr{L}$ is $4\sqrt{\underline{\lambda}(\mathbf{G}_a)}$. Therefore,

$$\mathbf{e}_\psi^{\mathsf{T}} \mathbf{e}_\psi \quad \leq \quad \frac{\sqrt{\underline{\lambda}(\mathbf{G}_a)}(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2}{4 \underline{\lambda}(\mathbf{G}_a)(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2} \psi,$$

in $\mathscr{L}$, and it follows that

$$b_1 \mathbf{e}_\psi^{\mathsf{T}} \mathbf{e}_\psi \leq \frac{b_1 \sqrt{\underline{\lambda}(\mathbf{G}_a)}(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2}{4 \underline{\lambda}(\mathbf{G}_a)(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2} \psi.$$

Setting

$$b_1 = \frac{4\sqrt{\underline{\lambda}(\mathbf{G}_a)}(c - 2\sqrt{\underline{\lambda}(\mathbf{G}_a)})^2}{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2},$$

gives $b_1\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi \leq \tilde{\psi}$. It remains to find b_2 such that $b_2\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi \geq \tilde{\psi}$. From (4.37), it follows that

$$\begin{aligned}
\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi &\geq \frac{\cos^2(\phi_e/2)\underline{\lambda}(\mathbf{G}_a)}{4(\underline{\lambda}(\mathbf{G}_a) - \tilde{E})}\tilde{E} \\
&= \frac{\cos^2(\phi_e/2)\underline{\lambda}(\mathbf{G}_a)}{4(\underline{\lambda}(\mathbf{G}_a) - \sin^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)}\tilde{E} \\
&\geq \frac{\cos^2(\phi_e/2)\underline{\lambda}(\mathbf{G}_a)}{4\underline{\lambda}(\mathbf{G}_a)(1 - \sin^2(\phi_e/2))}\tilde{E} \\
&= \tfrac{1}{4}\tilde{E} \\
&= \tfrac{1}{4}\tilde{\psi}\tilde{\delta}. \qquad (4.57)
\end{aligned}$$

The minimum value of $\tilde{\delta}$ in $\mathscr{L}$ is $2\sqrt{\underline{\lambda}(\mathbf{G}_a)}$. Therefore,

$$\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi \geq \tfrac{1}{2}\sqrt{\underline{\lambda}(\mathbf{G}_a)}\tilde{\psi}.$$

It follows that

$$b_2\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi \geq \tfrac{1}{2}b_2\sqrt{\underline{\lambda}(\mathbf{G}_a)}\tilde{\psi},$$

and setting

$$b_2 = \frac{2}{\sqrt{\underline{\lambda}(\mathbf{G}_a)}},$$

gives $b_2\mathbf{e}_\psi^\mathsf{T}\mathbf{e}_\psi \geq \tilde{\psi}$. □

Mentioned previously was the fact that for $\mathbf{G}_a = \mathbf{1}$, the error vector $\mathbf{e}_\psi$ is related to the quaternion representation of the attitude error $\tilde{\mathbf{R}}$, which is given by $\mathbf{q} = [\,\boldsymbol{\epsilon}^\mathsf{T}\ \eta\,]^\mathsf{T} \in \mathbb{S}^3$, where $\eta \in \mathbb{R}$ and $\boldsymbol{\epsilon} \in \mathbb{R}^3$ are the scalar and vector parts of $\mathbf{q}$, respectively. This can be shown in the following way. Recall that, in terms of the axis and angle of Euler's theorem, $\mathbf{e}_\gamma$ can be expressed by (4.34). Substituting $\mathbf{G}_a = \mathbf{1}$ into (4.34) gives $\mathbf{e}_\gamma = \sin(\phi_e)\mathbf{a}_e$. Further substituting $\mathbf{G}_a = \mathbf{1}$ into the expression for $\tilde{E}$ given by (4.24)

gives $\tilde{E} = \sin^2(\phi_e/2)$. Therefore, with $\mathbf{G}_a = \mathbf{1}$,

$$\begin{aligned}
\mathbf{e}_\psi &= \frac{\sin(\phi_e)}{2\sqrt{1-\sin^2\phi_e}}\mathbf{a}_e \\
&= \frac{\sin(\phi_e)}{2\cos(\phi_e/2)}\mathbf{a}_e \\
&= \frac{2\sin(\phi_e/2)\cos(\phi_e/2)}{2\cos(\phi_e/2)}\mathbf{a}_e \\
&= \sin(\phi_e/2)\mathbf{a}_e,
\end{aligned}$$

and $\mathbf{e}_\psi = \boldsymbol{\epsilon}$.

4.3.3 Gibbs Cost Function

Two cost functions on $SO(3)$, γ and ψ, have just been explored. Here, one more cost function on $SO(3)$ is introduced, which is denoted $\theta : SO(3) \times SO(3) \to \mathbb{R}^+$. This cost function is given by

$$\theta(\hat{\mathbf{R}}, \mathbf{R}) = \ln(\underline{\lambda}(\mathbf{G}_a)) - \ln(\underline{\lambda}(\mathbf{G}_a) - E(\hat{\mathbf{R}}, \mathbf{R})). \tag{4.58}$$

In terms of the group error $\tilde{\mathbf{R}}$, write $\theta(\hat{\mathbf{R}}, \mathbf{R}) = \tilde{\theta}(\tilde{\mathbf{R}})$, $\tilde{\theta} : SO(3) \to \mathbb{R}^+$, where

$$\tilde{\theta}(\tilde{\mathbf{R}}) = \ln(\underline{\lambda}(\mathbf{G}_a)) - \ln(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})). \tag{4.59}$$

As will be seen in Sec. 4.3.4, the use of θ as a cost function results in an even more aggressive error vector than both γ and ψ. The cost function is referred to as the Gibbs cost function since, when $\mathbf{G}_a = \mathbf{1}$, the resulting error vector is equivalent to the Gibbs representation of the attitude error $\tilde{\mathbf{R}}$. When $\mathbf{G}_a = \mathbf{1}$, the cost function also reduces to a popular error function, related to the Gibbs representation, used in the stability analysis of attitude control methods, see for example in [64]. As for ψ, the useful properties of θ are explored in the following proposition.

Proposition 4.3. *Consider the attitude error function $\tilde{\theta} : D \to \mathbb{R}^+$ given by*

$$\tilde{\theta}(\tilde{\mathbf{R}}) = \ln(\underline{\lambda}(\mathbf{G}_a)) - \ln(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})), \tag{4.60}$$

and associated attitude error vector

$$\mathbf{e}_\theta = \frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee, \tag{4.61}$$

where $D = \left\{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{E}(\tilde{\mathbf{R}}) \leq \underline{\lambda}(\mathbf{G}_a)\right\}$ *is the allowable domain of* $\tilde{\psi}$*. The attitude error vector,* $\mathbf{e}_\psi$*, is well defined in the sublevel set*

$$\mathscr{L} = \left\{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\theta}(\tilde{\mathbf{R}}) < \ln(\underline{\lambda}(\mathbf{G}_a))\right\}, \tag{4.62}$$

where $\mathscr{L} \subset D$ *and* $D \setminus \mathscr{L} = \{\tilde{\mathbf{R}} \in SO(3) \mid E(\tilde{\mathbf{R}}) = \underline{\lambda}(\mathbf{G}_a)\}$*. The attitude error function* $\tilde{\theta}$ *has the following properties.*

(i) $\tilde{\theta}$ *is positive definite about* $\tilde{\mathbf{R}} = \mathbf{1}$*.*

(ii) The gradient of $\tilde{\theta}$ *with respect to the Riemannian metric defined in* (4.2) *is* $\nabla\tilde{\theta}(\tilde{\mathbf{R}}) = \mathbf{e}_\theta^\times \tilde{\mathbf{R}}$*.*

(iii) The only critical point of $\tilde{\theta}$ *in* $\mathscr{L}$ *is* $\tilde{\mathbf{R}} = \mathbf{1}$*.*

(iv) Let $\Omega_c = \{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\theta}(\tilde{\mathbf{R}}) \leq c\}$*, where* $c < \frac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a))$*. Then, for all* $\tilde{\mathbf{R}} \in \Omega_c$ *there exists constants* b_1 *and* b_2 *such that*

$$b_1||\mathbf{e}_\theta||_2^2 \leq \tilde{\theta}(\tilde{\mathbf{R}}) \leq b_2||\mathbf{e}_\theta||_2^2, \tag{4.63}$$

where

$$b_1 = \frac{16\underline{\lambda}(\mathbf{G}_a)^2 e^{-2c}}{(\bar{\lambda}(\mathbf{G}_a) + \underline{\lambda}(\mathbf{G}_a))^2} \tag{4.64}$$

and

$$b_2 = \frac{4\underline{\lambda}(\mathbf{G}_a)c}{1 - e^{-c}}. \tag{4.65}$$

Proof The proof proceeds in the exact same manner as Proposition 4.2. Substituting (4.24) into (4.60) yields

$$\tilde{\theta}(\tilde{\mathbf{R}}) = \ln(\underline{\lambda}(\mathbf{G}_a)) - \ln(\underline{\lambda}(\mathbf{G}_a) - \sin^2(\phi_e/2)\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e). \tag{4.66}$$

Thus, it follows that $\tilde{\theta} \geq 0$ and $\tilde{\theta} = 0$ if and only if $\phi_e \in \{0,\ \pm 2\pi,\ \pm 4\pi, \ldots\}$, which implies that $\tilde{\mathbf{R}} = \mathbf{1}$. This proves (i).

To prove (ii), let $\mathbf{\Gamma}(\epsilon)$ be a curve at $\tilde{\mathbf{R}}$ given by $\mathbf{\Gamma}(\epsilon) = \exp(\epsilon\mathbf{S})\tilde{\mathbf{R}}$, where $\epsilon \in \mathbb{R}$ and $\mathbf{S} \in \mathfrak{so}(3)$. Then, the gradient of $\tilde{\psi}$ with respect to the Riemannian metric defined in (4.2) is defined as

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\theta}(\mathbf{\Gamma}(\epsilon))\bigg|_{\epsilon=0} = \langle\nabla\tilde{\theta}(\tilde{\mathbf{R}}), \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}. \tag{4.67}$$

From the definition of $\tilde{\theta}$ and $\mathbf{\Gamma}$ it follows that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\psi}(\mathbf{\Gamma}(\epsilon))\bigg|_{\epsilon=0} &= \frac{-1}{2(\underline{\lambda}(\mathbf{G}_a)-\tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a\mathbf{S}\tilde{\mathbf{R}}) && (4.68)\\
&= \frac{1}{2(\underline{\lambda}(\mathbf{G}_a)-\tilde{E}(\tilde{\mathbf{R}}))}\langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a), \mathbf{S}\rangle && (4.69)\\
&= \frac{1}{2(\underline{\lambda}(\mathbf{G}_a)-\tilde{E}(\tilde{\mathbf{R}}))}\langle \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}}, \mathbf{S}\tilde{\mathbf{R}}\rangle_{\tilde{\mathbf{R}}}, && (4.70)
\end{aligned}$$

which implies that

$$\begin{aligned}
\nabla\tilde{\theta}(\tilde{\mathbf{R}}) &= \frac{1}{2(\underline{\lambda}(\mathbf{G}_a)-\tilde{E}(\tilde{\mathbf{R}}))}\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)\tilde{\mathbf{R}} && (4.71)\\
&= \mathbf{e}_\theta^\times\tilde{\mathbf{R}}, && (4.72)
\end{aligned}$$

proving (ii).

To show that the only critical point in $\mathscr{L}$ is $\tilde{\mathbf{R}}=\mathbf{1}$, observe that all critical points must satisfy $\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)=\mathbf{0}$. Alternatively,

$$\tilde{\mathbf{R}}\mathbf{M}_a-\tilde{\mathbf{R}}^\mathsf{T}\mathbf{M}_a=\mathbf{0}.$$

The values of $\tilde{\mathbf{R}}$ that satisfy $\tilde{\mathbf{R}}\mathbf{M}_a-\tilde{\mathbf{R}}^\mathsf{T}\mathbf{M}_a=\mathbf{0}$ are $\tilde{\mathbf{R}}=\mathbf{1}$ and $\tilde{\mathbf{R}}=\exp(\pi\mathbf{v}^\times)$, where [17, p. 553]

(a) $\mathbf{v}\in\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ for distinct λ_1, λ_2, and λ_3,

(b) $\mathbf{v}\in\{\alpha\mathbf{v}_i+\beta\mathbf{v}_j \mid \alpha^2+\beta^2=1\}\cup\mathbf{v}_k$ for $i,j,k\in\{1,2,3\}$, $\lambda_i=\lambda_j\neq\lambda_k$, and

(c) $\mathbf{v}\in\mathbb{S}^2$ for $\lambda_1=\lambda_2=\lambda_3$.

Now, evaluating (4.24) at the critical points observe that $\tilde{E}(\mathbf{1})=0$, and

(a) $\tilde{E}(\exp(\pi\mathbf{v}_1^\times))=\mathbf{v}_1^\mathsf{T}\mathbf{G}_a\mathbf{v}_1=\underline{\lambda}(\mathbf{G}_a)$,

(b) $\tilde{E}(\exp(\pi\mathbf{v}^\times))=\underline{\lambda}(\mathbf{G}_a)$, where $\mathbf{v}\in\mathbb{S}^2$, and $\lambda_1=\lambda_2=\lambda_3$.

Turning to (4.60), the only critical point of ψ that lies in $\mathscr{L}$ is $\tilde{\mathbf{R}}=\mathbf{1}$.

For item (iv), note that

$$\mathbf{e}_\theta=\frac{1}{2(\underline{\lambda}(\mathbf{G}_a)-\tilde{E})}\mathbf{e}_\gamma, \tag{4.73}$$

and therefore from (4.36) and (4.37)

$$\frac{1}{4(\underline{\lambda}(\mathbf{G}_a)-\tilde{E})^2}\tilde{E} \leq \mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta \leq \frac{(\bar{\lambda}(\mathbf{G}_a)+\underline{\lambda}(\mathbf{G}_a))^2}{16\underline{\lambda}(\mathbf{G}_a)(\underline{\lambda}(\mathbf{G}_a)-\tilde{E})^2}\tilde{E}. \tag{4.74}$$

By assumption $\tilde{\mathbf{R}} \in \Omega_c$ and therefore from the definition of Ω_c and $\tilde{\theta}$ it follows that

$$\underline{\lambda}(\mathbf{G}_a)e^{-c} \leq \underline{\lambda}(\mathbf{G}_a) - E(\tilde{\mathbf{R}}) \leq \underline{\lambda}(\mathbf{G}_a) \tag{4.75}$$

for all $\tilde{\mathbf{R}} \in \Omega_c$. Therefore from (4.74),

$$\frac{1}{4\underline{\lambda}(\mathbf{G}_a)^2}\tilde{E} \leq \mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta \leq \frac{(\bar{\lambda}(\mathbf{G}_a)+\underline{\lambda}(\mathbf{G}_a))^2}{16\underline{\lambda}(\mathbf{G}_a)^3e^{-2c}}\tilde{E}. \tag{4.76}$$

For convenience, define constants α and β such that

$$\alpha\tilde{E} \leq \mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta \leq \beta\tilde{E}, \tag{4.77}$$

where

$$\alpha = \frac{1}{4\underline{\lambda}(\mathbf{G}_a)^2} \tag{4.78}$$

and

$$\beta = \frac{(\bar{\lambda}(\mathbf{G}_a)+\underline{\lambda}(\mathbf{G}_a))^2}{16\underline{\lambda}(\mathbf{G}_a)^3e^{-2c}}. \tag{4.79}$$

Recalling the definition of $\tilde{\theta}$,

$$\tilde{\theta} = \ln\left(\frac{\underline{\lambda}(\mathbf{G}_a)}{\underline{\lambda}(\mathbf{G}_a)-\tilde{E}}\right), \tag{4.80}$$

it follows that

$$\ln\left(\frac{\underline{\lambda}(\mathbf{G}_a)}{\underline{\lambda}(\mathbf{G}_a)-1/\beta\mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta}\right) \leq \tilde{\theta} \leq \ln\left(\frac{\underline{\lambda}(\mathbf{G}_a)}{\underline{\lambda}(\mathbf{G}_a)-1/\alpha\mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta}\right), \tag{4.81}$$

and therefore

$$\frac{\underline{\lambda}(\mathbf{G}_a)}{\underline{\lambda}(\mathbf{G}_a)-1/\beta\mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta} \leq e^\theta \leq \frac{\underline{\lambda}(\mathbf{G}_a)}{\underline{\lambda}(\mathbf{G}_a)-1/\alpha\mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta}. \tag{4.82}$$

From these two inequalities it can be shown that

$$\alpha\underline{\lambda}(\mathbf{G}_a)(1-e^{-\theta}) \leq \mathbf{e}_\theta^\mathsf{T}\mathbf{e}_\theta \leq \beta\underline{\lambda}(\mathbf{G}_a)(1-e^{-\theta}). \tag{4.83}$$

Now, b_1 and b_2 are sought such that

$$b_1\beta\underline{\lambda}(\mathbf{G}_a)(1-e^{-\theta}) \leq \theta \leq b_2\alpha\underline{\lambda}(\mathbf{G}_a)(1-e^{-\theta}). \tag{4.84}$$

Equation (4.84) is satisfied when $\tilde{\theta}(\tilde{\mathbf{R}}) = 0$ for any $0 < b_1 \leq b_2 < \infty$. Now, assume that $\tilde{\theta}(\tilde{\mathbf{R}}) \neq 0$ and continue from (4.84). Define $g(\tilde{\theta}) = \theta/(1-e^{-\theta})$, which is continuously differentiable on the domain $\Omega_c \setminus \{0\}$. Then, it follows that b_1 and b_2 must satisfy $b_1\beta\underline{\lambda}(\mathbf{G}_a) \leq g(\tilde{\theta})$ and $b_2\alpha\underline{\lambda}(\mathbf{G}_a) \geq g(\tilde{\theta})$. Posed in this way, constants b_1 and b_2 can be found as

$$b_1 = \frac{1}{\beta\underline{\lambda}(\mathbf{G}_a)} \inf_{\tilde{\theta}\neq 0} g(\tilde{\theta}) \tag{4.85}$$

and

$$b_2 = \frac{1}{\alpha\underline{\lambda}(\mathbf{G}_a)} \sup_{\tilde{\theta}\neq 0} g(\tilde{\theta}). \tag{4.86}$$

Starting with b_2, it can be shown that for all $\tilde{\theta} > 0$ $g(\tilde{\theta})$ is strictly increasing. Consequently, for all $\tilde{\theta} > 0$ it follows that $g(c) \geq g(\tilde{\theta})$. Thus, set $b_2 = g(c)/(\alpha\underline{\lambda}(\mathbf{G}_a))$. To find b_1 note that since $g(\tilde{\theta})$ is strictly decreasing for all $\tilde{\theta} > 0$ it follows that

$$b_1 = \frac{1}{\beta\underline{\lambda}(\mathbf{G}_a)} \lim_{\tilde{\theta}\to 0^+} g(\tilde{\theta}) = \frac{1}{\beta\underline{\lambda}(\mathbf{G}_a)}. \tag{4.87}$$

This proves (iv). □

Similar to the relationship between $\mathbf{e}_\psi$ and the quaternion, the error vector $\mathbf{e}_\theta$ is related to the Gibbs representation of $\tilde{\mathbf{R}}$, denoted $\mathbf{p} \triangleq \tan(\phi_e/2)\mathbf{a}_e$. This can be shown in the following way. Recall that, in terms of the axis and angle of Euler's theorem, $\mathbf{e}_\gamma$ can be expressed by (4.34). As in Sec. 4.3.2, substituting $\mathbf{G}_a = \mathbf{1}$ into (4.73) yields

$$\begin{aligned}
\mathbf{e}_\theta &= \frac{\sin(\phi_e)}{2(1-\sin^2\phi_e)}\mathbf{a}_e \\
&= \frac{\sin(\phi_e)}{2\cos^2(\phi_e/2)}\mathbf{a}_e \\
&= \frac{2\sin(\phi_e/2)\cos(\phi_e/2)}{2\cos^2(\phi_e/2)}\mathbf{a}_e \\
&= \tan(\phi_e/2)\mathbf{a}_e,
\end{aligned}$$

and therefore with $\mathbf{G}_a = \mathbf{1}$, $\mathbf{e}_\theta = \mathbf{p}$. It is clear then, in the case of $\mathbf{G}_a = \mathbf{1}$, the error vector $\mathbf{e}_\theta$ is

4.3.4 Cost Function and Error Vector Comparison

In the previous sections, it was shown that when $\mathbf{G}_a = \mathbf{1}$, the error vectors associated with ψ and θ are much more aggressive than the error vector associated with E when large estimation errors are present. This property, however, is true for any $\mathbf{G}_a > 0$. To observe this, first bounds on the norms of $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$ will be found and it will be shown that the lower bounds of $||\mathbf{e}_\psi||_2^2$ and $||\mathbf{e}_\theta||_2^2$ are monotonically increasing functions of ϕ_e while the upper bound of $||\mathbf{e}_\gamma||_2^2$ is not. First, recall from (4.36) that

$$\begin{aligned} ||\mathbf{e}_\gamma||_2^2 &\leq \sin^2(\phi_e/2)\cos^2(\phi_e/2)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2 + \tfrac{1}{4}\sin^2(\phi_e/2)(\bar{\lambda}(\mathbf{G}_a) - \underline{\lambda}(\mathbf{G}_a))^2 \\ &= \tfrac{1}{4}\sin^2(\phi_e)(\mathbf{a}_e^\mathsf{T}\mathbf{G}_a\mathbf{a}_e)^2 + \tfrac{1}{4}\sin^2(\phi_e/2)(\bar{\lambda}(\mathbf{G}_a) - \underline{\lambda}(\mathbf{G}_a))^2) \\ &\leq \tfrac{1}{4}\bar{\lambda}(\mathbf{G}_a)^2(\sin^2(\phi_e) + \sin^2(\phi_e/2)(r-1)^2), \end{aligned} \tag{4.88}$$

where $r = \underline{\lambda}(\mathbf{G}_a)/\bar{\lambda}(\mathbf{G}_a)$. It can be shown that, for any value of $r \in (0, 1]$, there exists a constant $a \in [\pi/2, \arccos(-1/4))$ such that the upper bound of $||\mathbf{e}_\gamma||_2^2$ is monotonically decreasing on the domain $\phi_e \in [a, \pi]$. In fact, if $r = 1$, then $||\mathbf{e}_\gamma||_2 \to 0$ as $\phi_e \to \pi$. To see this let $g(\phi_e) = \sin^2(\phi_e) + \sin^2(\phi_e/2)(r-1)^2$ and consider

$$\frac{\mathrm{d}}{\mathrm{d}\phi_e} g(\phi_e) = \sin(\phi_e/2)\cos(\phi_e/2)((r-1)^2 + 4\cos(\phi_e)). \tag{4.89}$$

This indicates that, on the domain $\phi_e \in [0, \pi]$, $g(\phi_e)$ has local extrema at $\phi_e = 0$, $\phi_e = \pi$, and $\phi_e = \arccos(-\frac{1}{4}(r-1)^2)$. For any value of r the extrema at $\arccos(-\frac{1}{4}(r-1)^2)$ lies between $\phi_e = 0$ and $\phi_e = \pi$. Given that both $\phi_e = 0$ and $\phi_e = \pi$ are local minimizers of $g(\phi_e)$ and that $\phi_e = \arccos(-\frac{1}{4}(r-1)^2)$ is a local maximizer, it follows that $g(\phi_e)$ is monotonically decreasing on the domain $\phi_e \in [\arccos(-\frac{1}{4}(r-1)^2), \pi]$. The ratio r lies in the range $(0, 1]$. Evaluating $\arccos(-\frac{1}{4}(r-1)^2)$ at these limits, and noting that $\arccos(-\frac{1}{4}(r-1)^2)$ is monotonically decreasing in r on the domain $r \in (0, 1]$, yields $\pi/2 < \arccos(-\frac{1}{4}(r-1)^2) \leq \arccos(-\frac{1}{4})$.

Next, consider $||\mathbf{e}_\psi||_2^2$ and $||\mathbf{e}_\theta||_2^2$. Recall from (4.57) that $||\mathbf{e}_\psi||_2^2 \geq \frac{1}{4}\tilde{E}$, which implies that

$$||\mathbf{e}_\psi||_2^2 \geq \tfrac{1}{4}\underline{\lambda}(\mathbf{G}_a)\sin^2(\phi_e/2). \tag{4.90}$$

Thus, the norm of $\mathbf{e}_\psi$ is bounded from below by a function of ϕ_e that is monotonically increasing on the domain $\phi_e \in [0, \pi]$. Similarly for $||\mathbf{e}_\theta||_2^2$, from (4.76)

$||\mathbf{e}_\theta||_2^2 \geq \tilde{E}/(4\underline{\lambda}(\mathbf{G}_a))^2$, which implies that

$$||\mathbf{e}_\theta||_2^2 \geq \sin^2(\phi_e/2)/(4\underline{\lambda}(\mathbf{G}_a)), \tag{4.91}$$

which is also monotonically increasing on $\phi_e \in [0, \pi]$. Consequently, the norm of $\mathbf{e}_\psi$ is also bounded from below by a function of ϕ_e that is monotonically increasing on the domain $\phi_e \in [0, \pi]$.

Comparing the behavior of (4.88), (4.90), and (4.91) in the domain $\phi_e \in [a, \pi]$, it is clear that $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ are much more aggressive than $\mathbf{e}_\gamma$ for large attitude estimation errors. Therefore, the use of $\mathbf{e}_\psi$ or $\mathbf{e}_\theta$ results in faster convergence when large estimation errors are present.

The behavior of $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$ as ϕ_e approaches π has been investigated. The behavior of $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$ when ϕ_e approaches zero will now be discussed. Linearizing (4.34), (4.53), and (4.73) about $\phi_e = 0$ gives

$$\mathbf{e}_\gamma \approx \tfrac{1}{2}\phi_e \mathbf{G}_a \mathbf{a}_e, \tag{4.92}$$

$$\mathbf{e}_\psi \approx \frac{\phi_e}{4\sqrt{\underline{\lambda}(\mathbf{G}_a)}} \mathbf{G}_a \mathbf{a}_e, \tag{4.93}$$

and

$$\mathbf{e}_\theta \approx \frac{\phi_e}{4\underline{\lambda}(\mathbf{G}_a)} \mathbf{G}_a \mathbf{a}_e, \tag{4.94}$$

respectively. Factors of $1/(2\sqrt{\underline{\lambda}(\mathbf{G}_a)})$ and $1/(2\underline{\lambda}(\mathbf{G}_a))$ are respectively present in (4.93) and (4.94) but not in (4.92). As a consequence, $\mathbf{e}_\psi$ has the same linearization about $\phi_e = 0$ as $2\sqrt{\underline{\lambda}(\mathbf{G}_a)}\mathbf{e}_\gamma$ and, similarly, $\mathbf{e}_\theta$ has the same linearization about $\phi_e = 0$ as $2\underline{\lambda}(\mathbf{G}_a)\mathbf{e}_\gamma$. This fact will be useful when comparing the performance of the gradient-based observers derived from the different cost functions, which will be considered in Section 4.5.

An analytical comparison of $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$ has been provided in this section. The results of the comparison can be summarized in graphical form as presented in Fig. 4.1. In this figure, $||\mathbf{e}_\psi||_2$, $2||\mathbf{e}_\theta||_2$, and $2||\mathbf{e}_\gamma||_2$ are plotted against ϕ_e, where $\mathbf{G}_a = \mathbf{1}$. In this case, $||\mathbf{e}_\gamma||_2 = \frac{1}{2}\sin(\phi_e)$, $2||\mathbf{e}_\psi||_2 = \sin(\phi_e/2)$, and $2||\mathbf{e}_\theta||_2 = \tan(\phi/2)$. From Fig. 4.1 it is clear that $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ are monotonically increasing in ϕ_e, while $\mathbf{e}_\gamma$ is not. It is also seen that $\sqrt{\underline{\lambda}(\mathbf{G}_a)}\mathbf{e}_\psi$, $\underline{\lambda}(\mathbf{G}_a)\mathbf{e}_\theta$, and $\mathbf{e}_\gamma$ share the same linearization about $\phi_e = 0$, a property that was observed from (4.92), (4.93), and (4.94).

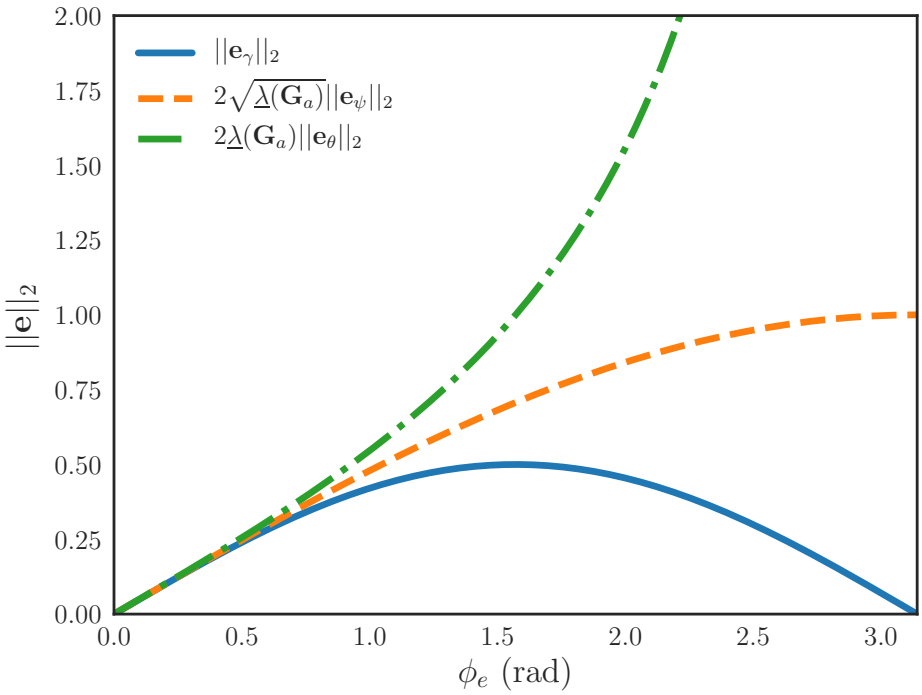


Figure 4.1: The error vectors associated with cost functions ψ and θ remain monotonically increasing with attitude estimation error ϕ_e, in stark contrast to the error vector associated with γ.

4.4 Gradient-Based Observer on SO(3)

4.4.1 Observer Design

The gradient-based observer for the Lie group $SO(3)$ is given by (3.19) and takes the form

$$\dot{\hat{\mathbf{R}}} = \hat{\mathbf{R}}\boldsymbol{\omega}_y^\times - k\mathbf{e}^\times\hat{\mathbf{R}}, \tag{4.95}$$

where $\mathbf{e}$ is the error vector associated with a cost function, $f : SO(3) \times SO(3) \to \mathbb{R}$, such that $\nabla_{\hat{\mathbf{R}}} f(\hat{\mathbf{R}}, \mathbf{R}) = \mathbf{e}^\times\hat{\mathbf{R}}$, and $k \in (0, \infty)$ is an observer gain. Several different cost functions given in (4.14), (4.40), and (4.60), that result in three distinct error vectors,

$$\mathbf{e}_\gamma = \boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee, \tag{4.96}$$

$$\mathbf{e}_\psi = \frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}}\boldsymbol{\mathcal{P}}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee, \tag{4.97}$$

and

$$\mathbf{e}_\theta = \frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))} \mathcal{P}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee, \tag{4.98}$$

have been discussed. However, in practice $\tilde{\mathbf{R}}$ is not available and therefore $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$ will need to be constructed from the attitude estimate $\hat{\mathbf{R}}$, the reference vectors $\mathbf{y}_a^j$, as well as the vector measurements $\mathbf{y}_b^j$. This is done by

$$\mathbf{e}_\gamma = \tfrac{1}{2} \textstyle\sum_{j=1}^n \mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j, \tag{4.99}$$

$$\mathbf{e}_\psi = \frac{1}{4\sqrt{\underline{\lambda}(\mathbf{G}_a) - \frac{1}{2}\sum_{j=1}^n k_j (1 - \mathbf{y}_a^{j\mathsf{T}} \hat{\mathbf{R}} \mathbf{y}_b^j)}} \textstyle\sum_{j=1}^n \mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j, \tag{4.100}$$

and

$$\mathbf{e}_\theta = \frac{1}{4(\underline{\lambda}(\mathbf{G}_a) - \frac{1}{2}\sum_{j=1}^n k_j (1 - \mathbf{y}_a^{j\mathsf{T}} \hat{\mathbf{R}} \mathbf{y}_b^j))} \textstyle\sum_{j=1}^n \mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j. \tag{4.101}$$

To see that (4.96), (4.97), and (4.98) are equivalent to (4.99), (4.100), and (4.101), consider

$$\begin{aligned}
\mathcal{P}_a(\tilde{\mathbf{R}}\mathbf{M}_a)^\vee &= \tfrac{1}{2}(\tilde{\mathbf{R}}\mathbf{M}_a - \mathbf{M}_a\tilde{\mathbf{R}}^\mathsf{T})^\vee \\
&= \tfrac{1}{2}(\tilde{\mathbf{R}} \textstyle\sum_{j=1}^n \mathbf{y}_a^j \mathbf{y}_a^{j\mathsf{T}} - \sum_{j=1}^n \mathbf{y}_a^j \mathbf{y}_a^{j\mathsf{T}} \tilde{\mathbf{R}}^\mathsf{T})^\vee \\
&= \tfrac{1}{2}(\textstyle\sum_{j=1}^n \hat{\mathbf{R}} \mathbf{y}_b^j \mathbf{y}_a^{j\mathsf{T}} - \sum_{j=1}^n \mathbf{y}_a^j \mathbf{y}_b^{j\mathsf{T}} \hat{\mathbf{R}}^\mathsf{T})^\vee \\
&= \tfrac{1}{2}(\textstyle\sum_{j=1}^n (\mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j)^\times)^\vee \\
&= \tfrac{1}{2} \textstyle\sum_{j=1}^n \mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j,
\end{aligned}$$

where (4.6) has been used. Next, from (4.13) $\tilde{E} = \frac{1}{2}\sum_{j=1}^n k_j (1 - \mathbf{y}_a^{j\mathsf{T}} \hat{\mathbf{R}} \mathbf{y}_b^j)$, which shows the equivalence.

Next, let us investigate the error dynamics of the observer. This will be useful in the Lyapunov analysis that will follow. From (4.4) and (4.95), it follows that the error in the attitude estimate, $\tilde{\mathbf{R}}$, evolves according to

$$\dot{\tilde{\mathbf{R}}} = \hat{\mathbf{R}}(\mathbf{b} + \boldsymbol{\mu})^\times \mathbf{R}^\mathsf{T} - k\mathbf{e}^\times \tilde{\mathbf{R}}. \tag{4.102}$$

4.4.2 Bias Estimation

In the case when the angular velocity measurement is corrupted by a constant, or slowly varying, bias the observer equations (4.95) can be augmented with a bias estimator. This is a popular technique employed in [10, 15, 16]. The attitude observer

with bias estimation is given by,

$$\dot{\hat{\mathbf{R}}} = \hat{\mathbf{R}}(\boldsymbol{\omega}_y - \hat{\mathbf{b}})^\times - \mathbf{e}^\times \hat{\mathbf{R}} \tag{4.103a}$$

$$\dot{\hat{\mathbf{b}}} = \mathbf{K}_b \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}, \tag{4.103b}$$

where $\hat{\mathbf{b}} \in \mathbb{R}^3$ is the estimate of $\mathbf{b}$, the constant bias present in (4.12), and $\mathbf{K}_b \in \mathbb{R}^{3\times 3}$ is a symmetric positive-definite matrix.

Again, it will be necessary for the stability analysis that follows to investigate the error dynamics associated with (4.103). Using (4.4) and (4.103), the error dynamics can be found to be

$$\dot{\tilde{\mathbf{R}}} = \hat{\mathbf{R}}(\tilde{\mathbf{b}} + \boldsymbol{\mu})^\times \mathbf{R}^\mathsf{T} - k\mathbf{e}^\times \tilde{\mathbf{R}} \tag{4.104a}$$

$$\dot{\tilde{\mathbf{b}}} = -\mathbf{K}_b \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}, \tag{4.104b}$$

where $\tilde{\mathbf{b}} = \mathbf{b} - \hat{\mathbf{b}}$ is the error in the estimate of $\mathbf{b}$.

4.4.3 Stability Results

In this section the stability properties of the gradient-based $SO(3)$ observers, given by both (4.95) and (4.103), with cost functions described in (4.40) and (4.60), will be explored. In particular, convergence of the state estimates to $\tilde{\mathbf{R}} = \mathbf{1}$ and $(\tilde{\mathbf{R}}, \tilde{\mathbf{b}}) = (\mathbf{1}, \mathbf{0})$, for (4.102) and (4.104) respectively, will be shown. It has been shown that global asymptotic stabilization on $SO(3)$ is not possible using continuous feedback [69]. As such, the stability results presented in this section are not global results. In fact, restrictions on the allowable set of initial conditions are required. This is to ensure that the error vectors $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ remain well defined. Similar restrictions have been considered previously in the literature [25, 63].

The stability properties of the gradient-based observer with cost function ψ are proved in Theorems 4.4 and 4.5, while the stability properties of the gradient-based observer with cost function θ are proved in Theorems 4.4 and 4.5.

Theorem 4.4. *Consider trajectories of* $\tilde{\mathbf{R}}$ *under* (4.102) *where* $\mathbf{e} = \mathbf{e}_\psi$. *Assume that there is no noise or bias associated with* $\boldsymbol{\omega}_y$, *that is* $\mathbf{b} = \boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\omega}_y = \boldsymbol{\omega}$. *Also assume that there is no noise associated with the vector measurements and therefore all* $\mathbf{y}_b^j$, $j = 1, \ldots, n$, *are measured exactly. If* $\tilde{\mathbf{R}}(0) \in \mathscr{L}$, *where* $\mathscr{L}$ *is defined in* (4.42), *then* $\mathbf{e}_\psi$ *exponentially approaches* $\mathbf{0}$ *and* $\tilde{\mathbf{R}}$ *asymptotically approaches* $\mathbf{1}$.

Proof Consider the following Lyapunov function candidate:

$$V(\tilde{\mathbf{R}}) = \tilde{\psi}(\tilde{\mathbf{R}}) = \sqrt{\underline{\lambda}(\mathbf{G}_a)} - \sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}. \tag{4.105}$$

Taking the derivative with respect to time,

$$\dot{V} = -\frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\mathbf{M}_a \dot{\tilde{\mathbf{R}}}).$$

From (4.8) and (4.102),

$$\begin{aligned} \dot{V} &= k\frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\mathbf{M}_a \mathbf{e}_\psi^\times \tilde{\mathbf{R}}) \\ &= k\frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\tilde{\mathbf{R}} \mathbf{M}_a \mathbf{e}_\psi^\times) \\ &= -2k||\mathbf{e}_\psi||_2^2. \end{aligned} \tag{4.106}$$

This implies $V(\tilde{\mathbf{R}}(t)) \leq V(\tilde{\mathbf{R}}(0))$ for all $t \geq 0$. Given $\tilde{\mathbf{R}}(0) \in \mathscr{L}$, it follows that $V(\tilde{\mathbf{R}}(t)) \leq V(\tilde{\mathbf{R}}(0)) < \sqrt{\underline{\lambda}(\mathbf{G})}$, which implies $\tilde{\mathbf{R}} \in \mathscr{L}$ $\forall t \geq 0$ and therefore $\mathbf{e}_\psi$ is bounded and well defined for all $t \geq 0$. Further, defining $\Omega_{V(\tilde{\mathbf{R}}(0))} = \{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\psi}(\tilde{\mathbf{R}}) \leq V(\tilde{\mathbf{R}}(0))\}$ it is seen that $\tilde{\mathbf{R}} \in \Omega_{V(\tilde{\mathbf{R}}(0))}$ for all $t \geq 0$ where by assumption $V(\tilde{\mathbf{R}}(0)) < \sqrt{\underline{\lambda}(\mathbf{G}_a)}$. Referring to Proposition 4.2, there exists two positive constants b_1 and b_2 such that

$$b_1||\mathbf{e}_\psi||_2^2 \leq V \leq b_2||\mathbf{e}_\psi||_2^2. \tag{4.107}$$

From Equations (4.106) and (4.107),

$$||\mathbf{e}_\psi(t)||_2 \leq \left(\frac{b_2}{b_1}\right)^{\frac{1}{2}} ||\mathbf{e}_\psi(0)||_2 \exp\left(-\frac{k}{2b_2}t\right), \tag{4.108}$$

and as a result $\mathbf{e}_\psi$ exponentially approaches $\mathbf{0}$. In $\mathscr{L}$, $\mathbf{e}_\psi = \mathbf{0}$ if and only if $\tilde{\mathbf{R}} = \mathbf{1}$, and therefore $\tilde{\mathbf{R}}$ asymptotically approaches $\mathbf{1}$. $\square$

The allowable set of initial conditions in Theorem 4.4, $\tilde{\mathbf{R}}(0) \in \mathscr{L}$, may be interpreted as an estimate of the region of attraction of $\tilde{\mathbf{R}} = \mathbf{1}$. As stated previously, if $\mathbf{y}_a^j$ and k_j are such that the eigenvalues of $\mathbf{M}_a$ are identical then $D = SO(3)$ and $D \setminus \mathscr{L}$ is given by the two-dimensional subset $\{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\mathbf{R}} = \exp(\pi \mathbf{v}^\times),\ \mathbf{v} \in \mathbb{S}^2\}$ where $\mathbf{e}_\psi$ is well defined in the sublevel set $\mathscr{L}$.

Theorem 4.5. *Consider trajectories of* $(\tilde{\mathbf{R}}, \tilde{\mathbf{b}})$ *under* (4.104)*, where* $\mathbf{e} = \mathbf{e}_\psi$*. Assume that there is no noise associated with* $\boldsymbol{\omega}_y$*, that is* $\boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\omega}_y = \boldsymbol{\omega} + \mathbf{b}$*. Also assume that there is no noise associated with the vector measurements, and therefore all* $\mathbf{y}_b^j$*,* $j = 1, \ldots, n$*, are measured exactly. If* $\tilde{\mathbf{R}}(0)$ *and* $\tilde{\mathbf{b}}(0)$ *are such that*

$$\tilde{\psi}(\tilde{\mathbf{R}}(0)) + \tilde{\mathbf{b}}(0)^\mathsf{T} \mathbf{K}_b^{-1} \tilde{\mathbf{b}}(0) < \sqrt{\underline{\lambda}(\mathbf{G}_a)}, \tag{4.109}$$

and $\boldsymbol{\omega}(t)$ *is bounded such that* $||\boldsymbol{\omega}(t)||_2 < \sigma$*,* $\sigma \in (0, \infty)$*,* $\forall t \geq 0$*, then* $(\tilde{\mathbf{R}}, \tilde{\mathbf{b}})$ *asymptotically approaches* $(\mathbf{1}, \mathbf{0})$*.*

Proof Consider the Lyapunov function candidate

$$V(\tilde{\mathbf{R}}, \tilde{\mathbf{b}}) = \sqrt{\underline{\lambda}(\mathbf{G}_a)} - \sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})} + \tilde{\mathbf{b}}^\mathsf{T} \mathbf{K}_b^{-1} \tilde{\mathbf{b}}.$$

Taking the derivative with respect to time,

$$\dot{V} = -\frac{1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\mathbf{M}_a \dot{\tilde{\mathbf{R}}}) + \tilde{\mathbf{b}}^\mathsf{T} \mathbf{K}_b^{-1} \dot{\tilde{\mathbf{b}}}.$$

From (4.8) and (4.104),

$$\begin{aligned}
\dot{V} &= \frac{-1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\mathbf{M}_a(\hat{\mathbf{R}} \tilde{\mathbf{b}}^\times \mathbf{R}^\mathsf{T} - k \mathbf{e}_\psi^\times \tilde{\mathbf{R}})) - 2 \tilde{\mathbf{b}}^\mathsf{T} \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}_\psi \\
&= \frac{-1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\mathbf{M}_a(\hat{\mathbf{R}} \tilde{\mathbf{b}}^\times \hat{\mathbf{R}}^\mathsf{T} \hat{\mathbf{R}} \mathbf{R}^\mathsf{T} - k \mathbf{e}_\psi^\times \tilde{\mathbf{R}})) - 2 \tilde{\mathbf{b}}^\mathsf{T} \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}_\psi \\
&= \frac{-1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \mathrm{tr}(\tilde{\mathbf{R}} \mathbf{M}_a(\hat{\mathbf{R}} \tilde{\mathbf{b}}^\times \hat{\mathbf{R}}^\mathsf{T} - k \mathbf{e}_\psi^\times)) - 2 \tilde{\mathbf{b}}^\mathsf{T} \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}_\psi \\
&= \frac{-1}{2\sqrt{\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})}} \Big(-k \mathrm{tr}(\tilde{\mathbf{R}} \mathbf{M}_a \mathbf{e}_\psi^\times) + \mathrm{tr}(\tilde{\mathbf{R}} \mathbf{M}_a (\hat{\mathbf{R}} \tilde{\mathbf{b}})^\times) \Big) - 2 \tilde{\mathbf{b}}^\mathsf{T} \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}_\psi \\
&= -2k ||\mathbf{e}_\psi||_2^2 + 2 \mathbf{e}_\psi^\mathsf{T} \hat{\mathbf{R}} \tilde{\mathbf{b}} - 2 \tilde{\mathbf{b}} \hat{\mathbf{R}}^\mathsf{T} \mathbf{e}_\psi \qquad (4.110) \\
&= -2k ||\mathbf{e}_\psi||_2^2. \qquad (4.111)
\end{aligned}$$

This implies $V(\tilde{\mathbf{R}}(t), \tilde{\mathbf{b}}(t)) \leq V(\tilde{\mathbf{R}}(0), \tilde{\mathbf{b}}(0))$ for all $t \geq 0$. Thus, $\tilde{\mathbf{b}}$ remains bounded for all $t \geq 0$. Referring to (4.109), $V(\tilde{\mathbf{R}}(t), \tilde{\mathbf{b}}(t)) \leq V(\tilde{\mathbf{R}}(0), \tilde{\mathbf{b}}(0)) < \sqrt{\underline{\lambda}(\mathbf{G}_a)}$, which implies $\tilde{\mathbf{R}} \in \mathscr{L}$ $\forall t \geq 0$ and therefore $\mathbf{e}_\psi$ is bounded and well defined for all $t \geq 0$. It can then be shown that $\ddot{V}$ is bounded for all $t \geq 0$. This implies $\dot{V}$ is uniformly continuous and by application of Barbalat's Lemma $\dot{V} \to 0$, and consequently $||\mathbf{e}_\psi||_2 \to 0$ as

$t \to \infty$. By property (iii) of Proposition 4.2, $||\mathbf{e}_\psi||_2 \to 0$ implies $\tilde{\mathbf{R}} \to \mathbf{1}$. Thus, $\tilde{\mathbf{R}}$ asymptotically approaches $\mathbf{1}$. Under the assumption that $\boldsymbol{\omega}$ remains bounded, it can be shown that $\ddot{\tilde{\mathbf{R}}}$ remains bounded. Thus, $\dot{\tilde{\mathbf{R}}}$ is uniformly continuous and $\dot{\tilde{\mathbf{R}}} \to \mathbf{0}$. From (4.104), this implies that $\hat{\mathbf{R}}\tilde{\mathbf{b}}^\times\mathbf{R}^\mathsf{T} \to \mathbf{0}$, which in turn implies that $\tilde{\mathbf{b}} \to \mathbf{0}$ as $t \to \infty$. $\square$

Theorem 4.6. *Consider trajectories of* $\tilde{\mathbf{R}}$ *under* (4.102) *where* $\mathbf{e} = \mathbf{e}_\theta$. *Assume that there is no noise or bias associated with* $\boldsymbol{\omega}_y$, *that is* $\mathbf{b} = \boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\omega}_y = \boldsymbol{\omega}$. *Also assume that there is no noise associated with the vector measurements and therefore all* $\mathbf{y}_b^j$, $j = 1, \ldots, n$, *are measured exactly. If* $\tilde{\mathbf{R}}(0) \in \mathscr{L}$, *where* $\mathscr{L}$ *is defined in* (4.62), *then* $\mathbf{e}_\theta$ *exponentially approaches* $\mathbf{0}$ *and* $\tilde{\mathbf{R}}$ *asymptotically approaches* $\mathbf{1}$.

Proof Consider the following Lyapunov function candidate:

$$V(\tilde{\mathbf{R}}) = \tilde{\theta}(\tilde{\mathbf{R}}) = \tfrac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a)) - \tfrac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})). \tag{4.112}$$

Taking the derivative with respect to time,

$$\dot{V} = -\frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a\dot{\tilde{\mathbf{R}}}).$$

From (4.8) and (4.102),

$$\begin{aligned}\dot{V} &= k\frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a\mathbf{e}_\theta^\times\tilde{\mathbf{R}}) \\ &= k\frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\tilde{\mathbf{R}}\mathbf{M}_a\mathbf{e}_\theta^\times) \\ &= -2k||\mathbf{e}_\theta||_2^2. \end{aligned} \tag{4.113}$$

This implies $V(\tilde{\mathbf{R}}(t)) \leq V(\tilde{\mathbf{R}}(0))$ for all $t \geq 0$. Given $\tilde{\mathbf{R}}(0) \in \mathscr{L}$, $V(\tilde{\mathbf{R}}(t)) \leq V(\tilde{\mathbf{R}}(0)) < \frac{1}{2}\ln(\underline{\lambda}(\mathbf{G}))$, which implies $\tilde{\mathbf{R}} \in \mathscr{L}$ $\forall t \geq 0$ and therefore $\mathbf{e}_\psi$ is bounded and well defined for all $t \geq 0$. Further, defining $\Omega_{V(\tilde{\mathbf{R}}(0))} = \{\tilde{\mathbf{R}} \in SO(3) \mid \tilde{\psi}(\tilde{\mathbf{R}}) \leq V(\tilde{\mathbf{R}}(0))\}$ it is seen that $\tilde{\mathbf{R}} \in \Omega_{V(\tilde{\mathbf{R}}(0))}$ for all $t \geq 0$ where by assumption $V(\tilde{\mathbf{R}}(0)) < \frac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a))$. Referring to Proposition 4.3, there exists two positive constants b_1 and b_2 such that

$$b_1||\mathbf{e}_\theta||_2^2 \leq V \leq b_2||\mathbf{e}_\theta||_2^2. \tag{4.114}$$

From Equations (4.113) and (4.114),

$$||\mathbf{e}_\theta(t)||_2 \leq \left(\frac{b_2}{b_1}\right)^{\frac{1}{2}} ||\mathbf{e}_\theta(0)||_2 \exp\left(-\frac{k}{2b_2}t\right), \tag{4.115}$$

and as a result $\mathbf{e}_\theta$ exponentially approaches $\mathbf{0}$. In $\mathscr{L}$, $\mathbf{e}_\theta = \mathbf{0}$ if and only if $\tilde{\mathbf{R}} = \mathbf{1}$, and therefore $\tilde{\mathbf{R}}$ asymptotically approaches $\mathbf{1}$. $\square$

Theorem 4.7. *Consider trajectories of* $(\tilde{\mathbf{R}}, \tilde{\mathbf{b}})$ *under* (4.104)*, where* $\mathbf{e} = \mathbf{e}_\theta$*. Assume that there is no noise associated with* $\boldsymbol{\omega}_y$*, that is* $\boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\omega}_y = \boldsymbol{\omega} + \mathbf{b}$*. Also assume that there is no noise associated with the vector measurements, and therefore all* $\mathbf{y}_b^j$*,* $j = 1, \ldots, n$*, are measured exactly. If* $\tilde{\mathbf{R}}(0)$ *and* $\tilde{\mathbf{b}}(0)$ *are such that*

$$\tilde{\theta}(\tilde{\mathbf{R}}(0)) + \tilde{\mathbf{b}}(0)^\mathsf{T}\mathbf{K}_b^{-1}\tilde{\mathbf{b}}(0) < \tfrac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a)), \tag{4.116}$$

and $\boldsymbol{\omega}(t)$ *is bounded such that* $||\boldsymbol{\omega}(t)||_2 < \sigma$*,* $\sigma \in (0, \infty)$*,* $\forall t \geq 0$*, then* $(\tilde{\mathbf{R}}, \tilde{\mathbf{b}})$ *asymptotically approaches* $(\mathbf{1}, \mathbf{0})$*.*

Proof Consider the Lyapunov function candidate

$$V(\tilde{\mathbf{R}}, \tilde{\mathbf{b}}) = \tfrac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a)) - \tfrac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}})) + \tilde{\mathbf{b}}^\mathsf{T}\mathbf{K}_b^{-1}\tilde{\mathbf{b}}.$$

Taking the derivative with respect to time,

$$\dot{V} = -\frac{1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a\dot{\tilde{\mathbf{R}}}) + 2\tilde{\mathbf{b}}^\mathsf{T}\mathbf{K}_b^{-1}\dot{\tilde{\mathbf{b}}}.$$

From (4.8) and (4.104),

$$\begin{aligned}
\dot{V} &= \frac{-1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a(\hat{\mathbf{R}}\tilde{\mathbf{b}}^\times\mathbf{R}^\mathsf{T} - k\mathbf{e}_\theta^\times\tilde{\mathbf{R}})) - 2\tilde{\mathbf{b}}^\mathsf{T}\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_\theta \\
&= \frac{-1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\mathbf{M}_a(\hat{\mathbf{R}}\tilde{\mathbf{b}}^\times\hat{\mathbf{R}}^\mathsf{T}\hat{\mathbf{R}}\mathbf{R}^\mathsf{T} - k\mathbf{e}_\theta^\times\tilde{\mathbf{R}})) - 2\tilde{\mathbf{b}}^\mathsf{T}\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_\psi \\
&= \frac{-1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\mathrm{tr}(\tilde{\mathbf{R}}\mathbf{M}_a(\hat{\mathbf{R}}\tilde{\mathbf{b}}^\times\hat{\mathbf{R}}^\mathsf{T} - k\mathbf{e}_\theta^\times)) - 2\tilde{\mathbf{b}}^\mathsf{T}\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_\theta \\
&= \frac{-1}{2(\underline{\lambda}(\mathbf{G}_a) - \tilde{E}(\tilde{\mathbf{R}}))}\left(-k\mathrm{tr}(\tilde{\mathbf{R}}\mathbf{M}_a\mathbf{e}_\theta^\times) + \mathrm{tr}(\tilde{\mathbf{R}}\mathbf{M}_a(\hat{\mathbf{R}}\tilde{\mathbf{b}})^\times)\right) - 2\tilde{\mathbf{b}}^\mathsf{T}\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_\theta \\
&= -2k||\mathbf{e}_\theta||_2^2 + 2\mathbf{e}_\theta^\mathsf{T}\hat{\mathbf{R}}\tilde{\mathbf{b}} - 2\tilde{\mathbf{b}}\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_\theta \qquad (4.117) \\
&= -2k||\mathbf{e}_\theta||_2^2. \qquad (4.118)
\end{aligned}$$

This implies $V(\tilde{\mathbf{R}}(t), \tilde{\mathbf{b}}(t)) \leq V(\tilde{\mathbf{R}}(0), \tilde{\mathbf{b}}(0))$ for all $t \geq 0$. Thus, $\tilde{\mathbf{b}}$ remains bounded for

all $t \geq 0$. Referring to (4.116), $V(\tilde{\mathbf{R}}(t), \tilde{\mathbf{b}}(t)) \leq V(\tilde{\mathbf{R}}(0), \tilde{\mathbf{b}}(0)) < \frac{1}{2}\ln(\underline{\lambda}(\mathbf{G}_a))$, which implies $\tilde{\mathbf{R}} \in \mathscr{L}\ \forall t \geq 0$ and therefore $\mathbf{e}_\theta$ is bounded and well defined for all $t \geq 0$. It can then be shown that $\ddot{V}$ is bounded for all $t \geq 0$. This implies $\dot{V}$ is uniformly continuous and by application of Barbalat's Lemma $\dot{V} \to 0$, and consequently $||\mathbf{e}_\psi||_2 \to 0$ as $t \to \infty$. By property (iii) of Proposition 4.2, $||\mathbf{e}_\psi||_2 \to 0$ implies $\tilde{\mathbf{R}} \to \mathbf{1}$. Thus, $\tilde{\mathbf{R}}$ asymptotically approaches $\mathbf{1}$. Under the assumption that $\boldsymbol{\omega}$ remains bounded, it can be shown that $\dddot{\tilde{\mathbf{R}}}$ remains bounded. Thus, $\ddot{\tilde{\mathbf{R}}}$ is uniformly continuous and $\dddot{\tilde{\mathbf{R}}} \to \mathbf{0}$. From (4.104), this implies that $\hat{\mathbf{R}}\tilde{\mathbf{b}}^\times\mathbf{R}^\mathsf{T} \to \mathbf{0}$, which in turn implies that $\tilde{\mathbf{b}} \to \mathbf{0}$ as $t \to \infty$. $\square$

4.5 Numerical Example

A numerical example of the attitude observers given in (4.95) and (4.103) is presented in this section. The attitude $\mathbf{R}$ is generated according to Poisson's equation, (4.4), where $\boldsymbol{\omega}(t) = -\pi^2/60(\cos(\pi/10t))\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^\mathsf{T}$ and $\mathbf{R}(0) = \exp(\phi(0)\mathbf{a}(0)^\times)$, where $\phi(0) = 0.99\pi$ (rad) and $\mathbf{a}(0) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$. It is assumed that the rigid body is equipped with sensors that measure $\mathbf{y}_b^j = \mathbf{R}_{\text{noise},j}^\mathsf{T}\mathbf{R}^\mathsf{T}\mathbf{y}_a^j$, $j = 1, 2$, where $\mathbf{R}_{\text{noise},j} \in SO(3)$ introduces noise into the measurement. The two reference vectors are $\mathbf{y}_a^1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\mathsf{T}$ and $\mathbf{y}_a^2 = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} & 0 \end{bmatrix}^\mathsf{T}$. The measurement noise is described by

$$\mathbf{R}_{\text{noise},j} = \exp(\mathbf{n}_j(t)^\times),$$

where $\mathbf{n}_j(t) = \begin{bmatrix} n_{j1}(t) & n_{j2}(t) & n_{j3}(t) \end{bmatrix}^\mathsf{T}$ with

$$\begin{aligned} n_{11}(t) &= 1/3^\circ \sin(10t) + 1/4^\circ \cos(12t), \\ n_{12}(t) &= 1/4^\circ \sin(15t) + 1/4^\circ \cos(10t), \\ n_{13}(t) &= 1/2^\circ \sin(18t) + 1/3^\circ \cos(20t), \end{aligned}$$

and

$$\begin{aligned} n_{21}(t) &= 1/2^\circ \sin(14t) + 1/3^\circ \cos(12t), \\ n_{22}(t) &= 1/4^\circ \sin(12t) + 1/3^\circ \cos(14t), \\ n_{23}(t) &= 1/2^\circ \sin(16t) + 1/4^\circ \cos(17t). \end{aligned}$$

When constructing $\mathbf{M}_a$ and $\mathbf{G}_a$, the gains are selected as $k_1 = 1$ and $k_2 = 0.5$, which gives

$$\mathbf{M}_a = \begin{bmatrix} 0.4 & 0.2 & 0 \\ 0.2 & 0.1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$\mathbf{G}_a = \begin{bmatrix} 1.1 & -0.2 & 0 \\ -0.2 & 1.4 & 0 \\ 0 & 0 & 0.5 \end{bmatrix},$$

with $\underline{\lambda}(\mathbf{G}_a) = 0.5$. The eigenvector associated with $\underline{\lambda}(\mathbf{G}_a)$ is $\mathbf{v}_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\mathsf{T}$.

The observers are initialized with $\hat{\mathbf{R}}(0) = \mathbf{1}$ and $\hat{\mathbf{b}} = \mathbf{0}$. This corresponds to an initial error in the yaw angle estimate of $-178.2°$, and initial pitch and roll estimates of $0°$. The large initial attitude estimation error is chosen to emphasize the convergence properties of the observers.

To make a fair comparison between the gradient-based observers with $\mathbf{e}_\gamma$, $\mathbf{e}_\psi$, and $\mathbf{e}_\theta$, it will be necessary to use different observer gains for each observer. This is due to the fact that the error vectors have different linearizations about $\phi_e = 0$, as was shown in Sec. 4.3.4. For $\mathbf{e}_\gamma$, the gains are selected as $\mathbf{K}_b = 0.1\mathbf{1}$ and $k = 0.5$, while for $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ the gains are scaled by a factor of $\sqrt{\underline{\lambda}(\mathbf{G}_a)}$ and $\underline{\lambda}(\mathbf{G}_a)$, respectively. In doing so, all error vectors have the same linearization about $\phi_e = 0$ and therefore it is expected that the different observers will have the same steady-state performance and a fair comparison can be made.

Two cases are considered. In the first case, the gradient-based observer without bias rejection, given by (4.95), is simulated. Here, the angular velocity measurement is assumed to be exact and $\mathbf{b} = \boldsymbol{\mu} = \mathbf{0}$. The results of the simulation are shown in Fig. 4.2 where the angle ϕ_e is plotted against time. The angle ϕ_e is extracted from $\tilde{\mathbf{R}}$ by the relation $\phi_e = \arccos(\frac{1}{2}(\mathrm{tr}(\tilde{\mathbf{R}}) - 1))$. Referring to Fig. 4.2, the attitude estimation error associated with the gradient-based observers derived from the costs γ, ψ, and θ approaches zero. However, as predicted the aggressive error vectors associated with ψ and θ result in faster convergence of the attitude estimate. The steady-state error of the observers, as shown in the inset of Fig. 4.2, remain equivalent. This is a property that was observed in Sec. 4.3.4.

In the second case, the gradient-based observer with bias rejection, given by (4.103), is simulated. It is assumed that the angular velocity measurement is corrupted by a constant bias of $\mathbf{b} = \begin{bmatrix} 0.01 & -0.01 & 0.01 \end{bmatrix}^\mathsf{T}$ (rad/s). Due to the fact

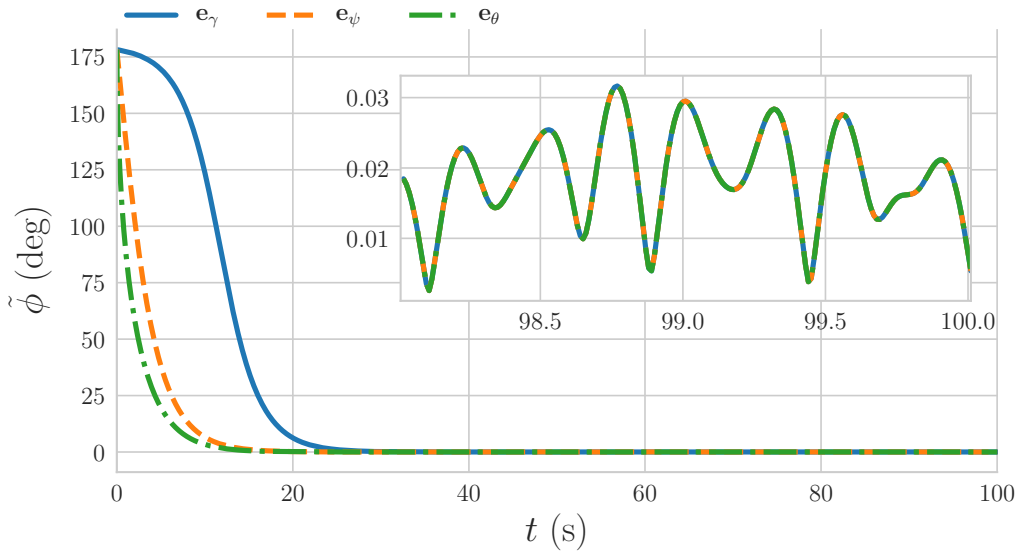


Figure 4.2: Time history of attitude error for the $SO(3)$ gradient-based observers. The last two seconds of the simulation are shown in the inset.

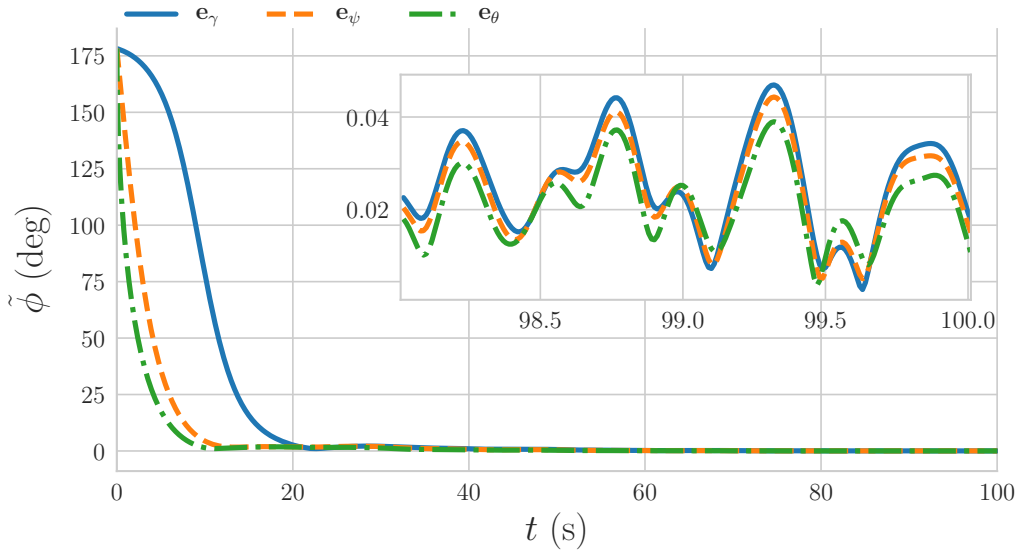


Figure 4.3: Time history of attitude error for the $SO(3)$ gradient-based observer with bias rejection. The last two seconds of the simulation are shown in the inset.

that the observer is initialized with large attitude estimation errors, a direct implementation of (4.103) will result in poor bias rejection. This is due to the fact that the bias estimate is the integral of the error vector. Consequently, integral windup will occur as the attitude estimate converges resulting in poor performance of the observer. As a solution, set $\dot{\hat{\mathbf{b}}} = \mathbf{0}$ if $||\mathbf{e}||_2 > 0.015$, and $\dot{\hat{\mathbf{b}}} = \mathbf{K}_b\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}$ if $||\mathbf{e}||_2 \leq 0.015$. Thus,

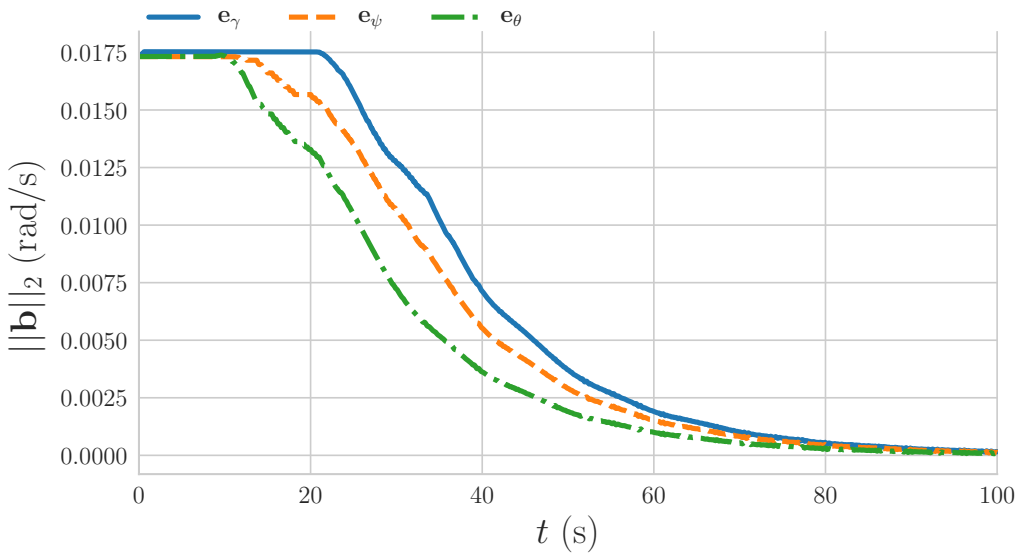


Figure 4.4: Time history of the bias estimation error norm.

integrator windup can be avoided and the effect of the attitude estimate transient on the bias estimate can be mitigated.

The results of the simulation are shown in Figs. 4.3 and 4.4. From Fig. 4.4, the bias estimates of all observers approach their true counterparts. However, due to the aggressive nature of $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$, the norms of $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ quickly approach the point where bias estimation can be implemented and integrator windup avoided. Due to the slow convergence associated with $\mathbf{e}_\gamma$, bias estimation is delayed resulting in slower convergence of the bias estimate.

4.6 Closing Remarks

In this chapter, three cost functions on $SO(3)$ and their corresponding gradient-based observers were explored. When coupled with the gradient-based observer design method discussed in Chapter 3, the first error function yields the well known nonlinear complementary filter on $SO(3)$ developed in [10]. Two alternate cost functions were then introduced two alternate that resulted in more aggressive innovation terms in the presence of large estimation errors. The introduction of these two cost functions are the main contributions of this chapter. A numerical simulation confirms the desirable convergence properties of the proposed attitude estimation methods. In practical terms, the proposed alternate cost functions lead to attitude observers that allow for faster convergence while maintaining equivalent steady-state performance when

compared to previously developed attitude observers.

Chapter 5

Simultaneous Localization and Mapping

5.1 Introduction

Simultaneous localization and mapping (SLAM) is a fundamental problem in robotics that is critical in enabling autonomous operation of ground, air, and marine vehicles. SLAM is the process of creating a map of the environment while, at the same time, estimating the pose of the vehicle relative to the map. Modern SLAM algorithms, such as FastSLAM [29], MonoSLAM [70], iSAM [31], GraphSLAM [30], and others [9], are posed in a probabilistic framework where a most likely estimate of the vehicle state and map is computed along with an associated uncertainty.

As an alternative to the probabilistic approach to SLAM, a nonlinear observer for the purpose of SLAM that evolves directly on the underlying Lie group is developed in this chapter. Several SLAM approaches that account for the inherent Lie group structure of the problem exist in the literature. For example, a scan matching SLAM observer is considered in [71]. Further, preintegration on a manifold is considered in [72], where the developed theory is used in conjunction with iSAM2 [32]. Error propagation, which is needed in probabilistic approaches, is considered directly on the special Euclidean group in [73]. Further, a geometric approach is also taken in [74] where the invariant EKF, a method developed in [75], is applied to the SLAM problem. In this chapter, a SLAM algorithm is developed based on the foundations presented in [14], which is similar to a recently developed SLAM algorithm proposed in [76]. The proposed SLAM algorithm involves propagating the pose and map estimates using angular and translational velocity measurements as well as vectorial measurements

of map features. The observer design is based on the theoretical foundations of [14], [45], and [16]. Thus, the proposed observer evolves directly on the underlying Lie group and therefore accounts for nonlinearities inherent in the three dimensional SLAM problem. Consequently, the proposed observer belongs to the same class of nonlinear observers developed for attitude estimation [10, 22, 24, 62, 65] and pose estimation [26, 27]. As in [14, 16, 45], the proposed observer can be decomposed into two components, a process model that mimics the kinematic equations of a rigid body as well as an innovation term that serves to drive the state estimates to their desired trajectories. This is similar to the scan matching SLAM observer presented in [71], which is based on the nonlinear observer design methods presented in [11] and [13]. In [71], the innovation term for the pose observer is related to the error between the pose estimate and the pose found from an Iterative Closest Point algorithm. In this paper, the innovation term is chosen based on the descent direction of a particular error function and is constructed using the vectorial measurements directly. The proposed method is similar to the pose observers presented in [26] and [27] and, in fact, builds upon these results by deriving a gradient-based observer for the full SLAM problem, rather than the pose estimation problem. As in [16] and [27], the problem of bias in the velocity measurements is also considered.

An independently developed SLAM algorithm, based on similar principles to the method given in this chapter, is presented in [76]. The SLAM observer in [76] and the method presented here are similar and share therefore share similar properties. They are both posed in a geometric framework, both observers evolve on the underlying SLAM Lie group, and they both have provable stability properties. The observer innovation terms introduced in [76] are derived from a Lyapunov stability analysis, while the innovation terms proposed in this chapter are derived from the gradient-based observer [14] discussed in Chapter 3. While the two observers share many similarities, they do differ in a number of ways. First, the innovation terms differ, which will result in distinct behavior of the two observers. Second, the proposed observer includes a gain mapping that allows for cross-coupling between pose and landmark innovations. Finally, the problem of bias estimation is explicitly handled in this chapter.

The main contribution of this chapter is the introduction of a gradient-based observer for the purposes of simultaneous localization and mapping. Further contributions are the proofs of convergence given in Theorems 5.3 and 5.4. The SLAM problem is discussed in Sec. 5.2 and the gradient-based observer for SLAM is proposed in 5.3. The stability properties of the proposed algorithm are proved mathematically

in Sec. 5.6 and demonstrated in simulation in Sec. 5.7. Finally, the algorithm is implemented in experiment in Sec. 5.8 and closing remarks are given in Sec. 5.9.

5.2 Problem Formulation

5.2.1 Rigid-Body Pose Kinematics

Consider a vehicle moving in three dimensional space. Let $\mathcal{F}_b$ denote the body-fixed frame of the vehicle and let $\mathcal{F}_a$ be an 'absolute' reference frame relative to which the attitude and position of the vehicle will be measured. Recall from Sec. 4.2 that the attitude of the vehicle is uniquely and globally described by the direction cosine matrix $\mathbf{C}_{ab} \in SO(3)$, where $\mathbf{C}_{ab}$ is the change of basis matrix mapping the coordinates of a vector resolved in $\mathcal{F}_b$ to $\mathcal{F}_a$. The position of the vehicle is described by $\underrightarrow{r}^{ba}$, where $\underrightarrow{r}^{ba}$ is the position of the origin of $\mathcal{F}_b$ relative to the origin of $\mathcal{F}_a$. The components of $\underrightarrow{r}^{ba}$ resolved in $\mathcal{F}_a$ are denoted $\mathbf{r}_a^{ba} \in \mathbb{R}^3$. Combining $\mathbf{C}_{ab}$ and $\mathbf{r}_a^{ba}$, the pose (i.e., the position and attitude) of the vehicle can be represented by the matrix $\mathbf{T}_{ab} \in SE(3)$, where the matrix Lie Group $SE(3)$ is the special Euclidean group given by

$$SE(3) = \left\{ \begin{bmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4} \mid \mathbf{C} \in SO(3), \mathbf{r} \in \mathbb{R}^3 \right\}, \tag{5.1}$$

and [17]

$$\mathbf{T}_{ab} = \begin{bmatrix} \mathbf{C}_{ab} & \mathbf{p}_a^{ba} \\ \mathbf{0} & 1 \end{bmatrix}. \tag{5.2}$$

The angular and translational velocities of the vehicle, resolved in $\mathcal{F}_b$, are denoted $\boldsymbol{\omega}_b^{ba} \in \mathbb{R}^3$ and $\mathbf{v}_b^{ba} \in \mathbb{R}^3$, respectively. The kinematic relationship that describes the evolution of $\mathbf{T}_{ab}$ through time is given by

$$\dot{\mathbf{T}}_{ab} = \mathbf{T}_{ab}\mathbf{A}_b, \tag{5.3}$$

where

$$\mathbf{A}_b = \begin{bmatrix} \boldsymbol{\omega}_b^{ba\times} & \mathbf{v}_b^{ba} \\ \mathbf{0} & 0 \end{bmatrix} \in \mathfrak{se}(3), \tag{5.4}$$

where $\mathfrak{se}(3)$ is the matrix Lie algebra of $SE(3)$. The matrix Lie algebra $\mathfrak{se}(3)$ is given by

$$\mathfrak{se}(3) = \left\{ \begin{bmatrix} \mathbf{S} & \mathbf{u} \\ \mathbf{0} & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4} \mid \mathbf{S} \in \mathfrak{so}(3), \mathbf{u} \in \mathbb{R}^3 \right\}. \tag{5.5}$$

As in 4.2, the standard matrix inner product is given by $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{tr}(\mathbf{X}^\mathsf{T}\mathbf{Y}),\ \ \forall \mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times n}$. Let $\boldsymbol{\mathcal{P}} : \mathbb{R}^{4\times 4} \to \mathfrak{se}(3)$ be the projection of $\mathbb{R}^{4\times 4}$ to $\mathfrak{se}(3)$. In particular,

$$\boldsymbol{\mathcal{P}}(\mathbf{M}) = \begin{bmatrix} \boldsymbol{\mathcal{P}}_a(\mathbf{M}_1) & \mathbf{m}_2 \\ \mathbf{0} & 0 \end{bmatrix}, \tag{5.6}$$

for all

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{m}_2 \\ \mathbf{m}_3^\mathsf{T} & m_4 \end{bmatrix} \in \mathbb{R}^{4\times 4}, \tag{5.7}$$

where $\mathbf{M}_1 \in \mathbb{R}^{3\times 3}$, $\mathbf{m}_2, \mathbf{m}_3 \in \mathbb{R}^3$, $m_4 \in \mathbb{R}$, and $\boldsymbol{\mathcal{P}}_a(\cdot)$ is the anti-symmetric projection operator defined in Sec. 4.2. It follows that, $\forall \mathbf{A} \in \mathfrak{se}(3)$, $\langle \mathbf{A}, \mathbf{M} \rangle = \langle \mathbf{A}, \boldsymbol{\mathcal{P}}(\mathbf{M}) \rangle$ [26]. Similarly, let $\boldsymbol{p} : \mathbb{R}^4 \to \mathcal{M}_0$ be the projection of $\mathbb{R}^4$ to $\mathcal{M}_0$ such that $\boldsymbol{p}(\mathbf{x}) = [\ x_1\ x_2\ x_3\ 0]^\mathsf{T}$ for all $\mathbf{x} = [\ x_1\ x_2\ x_3\ x_4]^\mathsf{T}$. It follows that $\mathbf{y}^\mathsf{T}\mathbf{x} = \mathbf{y}^\mathsf{T}\boldsymbol{p}(\mathbf{x})$, $\forall \mathbf{y} \in \mathcal{M}_0$ and $\mathbf{x} \in \mathbb{R}^4$.

The identities

$$\begin{aligned} \boldsymbol{\mathcal{P}}(\mathbf{T}^{-\mathsf{T}}\mathbf{x}\mathbf{y}^\mathsf{T}) &= \boldsymbol{\mathcal{P}}(\mathbf{T}\mathbf{x}\mathbf{y}^\mathsf{T}), & (5.8)\\ \boldsymbol{\mathcal{P}}((\mathbf{y}_1 - \mathbf{y}_2)\mathbf{y}_1^\mathsf{T}) &= \boldsymbol{\mathcal{P}}((\mathbf{y}_1 - \mathbf{y}_2)\mathbf{y}_2^\mathsf{T}), & (5.9)\\ ||\mathbf{T}\mathbf{x}||_2^2 &= ||\mathbf{x}||_2^2, & (5.10)\\ \boldsymbol{p}(\mathbf{T}^\mathsf{T}\mathbf{T}\mathbf{x}) &= \mathbf{x}, & (5.11) \end{aligned}$$

where $\mathbf{y}, \mathbf{y}_1, \mathbf{y}_2 \in \mathcal{M}_1$, $\mathbf{x} \in \mathcal{M}_0$, and $\mathbf{T} \in SE(3)$ [26], will be used in forthcoming derivations.

5.2.2 The SLAM Problem

Suppose there exists $\ell \in \mathbb{N}$, $\ell > 0$, point features in the environment surrounding the vehicle. The point features are referred to individually as landmarks and as the map when considered all together. The position of an arbitrary landmark, denoted point j, $j \in \{1, 2, \ldots, \ell\}$, may be measured relative to any reference frame. For example, an observer in, and at the origins of, $\mathcal{F}_b$ or $\mathcal{F}_a$ measures the position of point j as $\mathbf{r}_b^{jb} \in \mathbb{R}^3$ or $\mathbf{r}_a^{ja} \in \mathbb{R}^3$, respectively. The relationship between $\mathbf{r}_a^{ja}$ and $\mathbf{r}_b^{jb}$ is given by $\mathbf{r}_a^{ja} = \mathbf{C}_{ab}\mathbf{r}_b^{jb} + \mathbf{r}_a^{ba}$ [17]. As in [26], define $\mathcal{M}_1$ and $\mathcal{M}_0$ as submanifolds of $\mathbb{R}^4$ such that

$$\mathcal{M}_0 = \{\mathbf{x} = [\ x_1\ x_2\ x_3\ x_4\ ]^\mathsf{T} \in \mathbb{R}^4 \mid x_4 = 0\}, \tag{5.12}$$

and

$$\mathcal{M}_1 = \{\mathbf{x} = [\ x_1\ x_2\ x_3\ x_4\ ]^\mathsf{T} \in \mathbb{R}^4 \mid x_4 = 1\}. \tag{5.13}$$

Then, letting $\mathbf{m}_a^j = [\ \mathbf{r}_a^{ja\mathsf{T}}\ 1\ ]^\mathsf{T} \in \mathcal{M}_1$ and $\mathbf{m}_b^j = [\ \mathbf{r}_b^{jb\mathsf{T}}\ 1\ ]^\mathsf{T} \in \mathcal{M}_1$, the relationship between $\mathbf{m}_a^j$ and $\mathbf{m}_b^j$ is given by $\mathbf{m}_a^j = \mathbf{T}_{ab}\mathbf{m}_b^j$.

SLAM is the process of estimating the pose $\mathbf{T}_{ab}$ of the vehicle relative to the map $\{\mathbf{m}_a^1, \ldots, \mathbf{m}_a^\ell\}$ while simultaneously estimating the map. Ideally, it is desired to build an estimator such that the estimate of the map and pose approach their true counterparts. However, it is well known that the SLAM problem is unobservable, and consequently, the best that can be achieved is an estimate of the pose and map up to a constant pose transformation [77]. The SLAM problem considered in this chapter is stated as follows. Consider the system

$$\dot{\mathbf{T}}_{ab} = \mathbf{T}_{ab}\mathbf{A}_b, \tag{5.14a}$$

$$\dot{\mathbf{m}}_a^j = \mathbf{0}, \tag{5.14b}$$

and let $\mathbf{y}_j = \mathbf{T}_{ab}^{-1}\mathbf{m}_a^j \in \mathcal{M}_1$ be the system outputs. Given measurements of angular and translational velocity $\mathbf{A}_b$ as well as position measurements of stationary landmarks in the environment $\mathbf{y}_j$, design an observer such that the pose and map estimates, denoted respectively by $\mathbf{T}_{ae} \in SE(3)$ and $\hat{\mathbf{m}}_a^j \in \mathcal{M}_1$, satisfy $\mathbf{T}_{ae} \to \mathbf{T}\mathbf{T}_{ab}$ and $\hat{\mathbf{m}}_a^j \to \mathbf{T}\mathbf{m}_a^j$ as $t \to \infty$, for some constant pose transformation $\mathbf{T} \in SE(3)$.

It is assumed throughout this chapter that the identity of landmarks are known upon measurement and between measurements. Consequently, the data association problem is not treated. It is also assumed that landmark positions remained fixed with respect to frame $\mathcal{F}_a$. While the problems of data association and time-varying landmarks are important, they are outside the scope of this dissertation.

5.3 Gradient-Based Observer SLAM

The goal of this section is to apply the gradient-based observer design method discussed in Chapter 3 to the simultaneous localization and mapping problem. To do this, first the Lie group under which the state evolves must first be identified. Then, a Riemannian metric for the Lie group may be defined, which allows for the definition of the gradient of a function on the Lie group. Finally, a cost function may be defined and used in the gradient-based observer (3.19). The underlying Lie group of the SLAM problem is now discussed.

5.4 The SLAM Lie Group

First consider the manifold $\mathcal{M}_1$ and let $\mathbf{m}_x, \mathbf{m}_y \in \mathcal{M}_1$. As in [26], associate with every $\mathbf{m} \in \mathcal{M}_1$ an element of $\mathbb{R}^3$, denoted $\underline{\mathbf{m}} \in \mathbb{R}^3$, such that $\mathbf{m} = \begin{bmatrix} \underline{\mathbf{m}}^\mathsf{T} & 1 \end{bmatrix}^\mathsf{T}$. The manifold $\mathcal{M}_1$ is a Lie group with the group multiplication map given by

$$(\mathbf{m}_x, \mathbf{m}_y) \mapsto \begin{bmatrix} \underline{\mathbf{m}}_x + \underline{\mathbf{m}}_y \\ 1 \end{bmatrix}, \tag{5.15}$$

and inverse map

$$\mathbf{m}_x \mapsto \begin{bmatrix} -\underline{\mathbf{m}}_x \\ 1 \end{bmatrix}. \tag{5.16}$$

Next, consider the manifold $\mathcal{M}_1^\ell = \mathcal{M}_1 \times \ldots \mathcal{M}_1$, the Cartesian product of ℓ copies of $\mathcal{M}_1$. Since the Cartesian product of Lie groups is itself a Lie group [78], the manifold $\mathcal{M}_1^\ell$ is a Lie group under the group multiplication map

$$(m_x, m_y) \mapsto \left(\begin{bmatrix} \underline{\mathbf{m}}_{x1} + \underline{\mathbf{m}}_{y1} \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} \underline{\mathbf{m}}_{x\ell} + \underline{\mathbf{m}}_{y\ell} \\ 1 \end{bmatrix} \right), \tag{5.17}$$

and inverse map

$$m_x \mapsto \left(\begin{bmatrix} -\underline{\mathbf{m}}_{x1} \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} -\underline{\mathbf{m}}_{x\ell} \\ 1 \end{bmatrix} \right), \tag{5.18}$$

where $m_x, m_y \in \mathcal{M}_1^\ell$ with $m_x = (\mathbf{m}_{x1}, \ldots, \mathbf{m}_{x\ell})$ and $m_y = (\mathbf{m}_{y1}, \ldots, \mathbf{m}_{y\ell})$.

Consider the Cartesian product of $SE(3)$ and $\mathcal{M}_1^\ell$, $G = SE(3) \times \mathcal{M}_1^\ell$, where again G, the Cartesian product of two Lie groups, is a Lie group [78]. Given two elements of G, $g = (\mathbf{T}_g, m_g) \in G$ and $h = (\mathbf{T}_h, m_h) \in G$, the group multiplication operator $\star$ is defined as

$$g \star h = \left(\mathbf{T}_g \mathbf{T}_h, \begin{bmatrix} \underline{\mathbf{m}}_{g1} + \underline{\mathbf{m}}_{h1} \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} \underline{\mathbf{m}}_{g\ell} + \underline{\mathbf{m}}_{h\ell} \\ 1 \end{bmatrix} \right), \tag{5.19}$$

and the inversion operator $(\cdot)^{-1}$ is defined as

$$g^{-1} = \left(\mathbf{T}_g^{-1}, \begin{bmatrix} -\underline{\mathbf{m}}_{g1} \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} -\underline{\mathbf{m}}_{g\ell} \\ 1 \end{bmatrix} \right). \tag{5.20}$$

The tangent space of G at a point $x = (\mathbf{T}, m) \in G$ is $T_x G = T_{\mathbf{T}} SE(3) \times \mathcal{M}_0^\ell$, where $\mathcal{M}_0^\ell = \mathcal{M}_0 \times \mathcal{M}_0 \times \ldots \times \mathcal{M}_0$ and $T_{\mathbf{T}} SE(3)$ is the tangent space of $SE(3)$ at point $\mathbf{T} \in SE(3)$. The Lie algebra of G is the tangent space at the identity

element of G denoted $\mathfrak{g} = \mathfrak{se}(3) \times \mathcal{M}_0^\ell$. An element of $\mathcal{M}_0^\ell$, $a \in \mathcal{M}_0^\ell$, is expressed as $a = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_\ell)$, where $\mathbf{a}_j \in \mathcal{M}_0$ for all $j \in \{1, 2, \ldots, \ell\}$ [79].

The SLAM Lie group is now identified as the Lie group G, where the system state associated with (5.14) is given by $x = (\mathbf{T}_{ab}, m) \in G$, where $m = (\mathbf{m}_a^1, \ldots, \mathbf{m}_a^\ell) \in \mathcal{M}_1^\ell$. In terms of x, the system (5.14) can be expressed as $\dot{x} = (\mathbf{T}_{ab}\mathbf{A}_b, 0_{\mathcal{M}_0^\ell})$, where $(\mathbf{T}_{ab}\mathbf{A}_b, 0_{\mathcal{M}_0^\ell}) \in T_xG$ and $0_{\mathcal{M}_0^\ell} = (\mathbf{0}, \ldots, \mathbf{0}) \in \mathcal{M}_0^\ell$. For simplicity in notation, the subscript of $0_{\mathcal{M}_0^\ell}$ is dropped and the state equations are written as

$$\dot{x} = (\mathbf{T}_{ab}\mathbf{A}_b, 0). \tag{5.21}$$

The estimate of the state x is denoted $\hat{x} = (\mathbf{T}_{ae}, \hat{m}) \in G$, where $\mathbf{T}_{ae} \in SE(3)$ is the pose estimate and $\hat{m} = (\hat{\mathbf{m}}_a^1, \ldots, \hat{\mathbf{m}}_a^\ell) \in \mathcal{M}_1^\ell$ is the map estimate.

5.4.1 A Riemannian Metric on the SLAM Group

To find a Riemannian metric on the SLAM group, the approach of [17] is taken where first an inner product on the Lie algebra is first found. An inner product on $\mathfrak{g}$, denoted by $\langle \cdot, \cdot \rangle_\mathfrak{g} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, can be defined by

$$\langle (\mathbf{A}, a), (\mathbf{B}, b) \rangle_\mathfrak{g} \;=\; \langle \mathbf{A}, \mathbf{B} \rangle + \textstyle\sum_{j=1}^{\ell} \mathbf{a}_j^\mathsf{T} \mathbf{b}_j, \tag{5.22}$$

for all $\mathbf{A}, \mathbf{B} \in \mathfrak{se}(3)$ and $a, b \in \mathcal{M}_0^\ell$. Then, the inner product on $\mathfrak{g}$ determines a right-invariant Riemannian metric, $\langle \cdot, \cdot \rangle_g : T_gG \times T_gG \to \mathbb{R}$, by the relationship [17]

$$\langle v, u \rangle_g \;=\; \langle \mathbf{V}, \mathbf{U} \rangle_{\mathbf{T}_g} + \langle a, b \rangle_{m_g}, \tag{5.23}$$

for all $v = (\mathbf{V}, a) \in T_gG$, $u = (\mathbf{U}, b) \in T_gG$, and $g = (\mathbf{T}_g, m_g) \in G$, where $\langle \mathbf{V}, \mathbf{U} \rangle_{\mathbf{T}_g} \triangleq \langle \mathbf{V}\mathbf{T}_g^{-1}, \mathbf{U}\mathbf{T}_g^{-1} \rangle$ and $\langle a, b \rangle_{m_g} \triangleq \sum_{j=1}^{\ell} \mathbf{a}_j^\mathsf{T} \mathbf{b}_j$.

5.4.2 Cost Function Selection

Let $E : G \times G \to \mathbb{R}^+$ denote the error function that is sought. In [26], nonlinear observer design on $SE(3)$ is considered. In that paper the authors employ the cost function given by

$$E_{SE(3)}(\mathbf{T}_{ae}, \mathbf{T}_{ab}) = \tfrac{1}{2} \textstyle\sum_{j=1}^{\ell} ||\mathbf{T}_{ae}^{-1}\mathbf{m}_a^j - \mathbf{T}_{ab}^{-1}\mathbf{m}_a^j||_2^2, \tag{5.24}$$

which, similar to the $SO(3)$ cost function in Section 4.3.1, is related to the sum of the errors between predicted and true landmark observations. Following [26], in this

chapter the following error function is considered,

$$E(\hat{x}, x) \;=\; \tfrac{1}{2}\sum_{j=1}^{\ell} ||\mathbf{T}_{ae}^{-1}\hat{\mathbf{m}}_a^j - \mathbf{T}_{ab}^{-1}\mathbf{m}_a^j||_2^2, \tag{5.25}$$

which is the sum of the errors between predicted and true landmark observations. This error function has as similar form to the Lyapunov function used in [76]. It is important to note that $E(\hat{x}, x) = 0$ does not imply that $\hat{x} = x$.

From Definition 2.24, the gradient of E with respect to $\hat{x}$ is the unique tangent vector $\nabla_{\hat{x}} E(\hat{x}, x) \in T_{\hat{x}}G$ such that

$$d_1 E(\hat{x}, x) \circ v = \langle \nabla_{\hat{x}} E(\hat{x}, x), v \rangle_{\hat{x}}, \;\; \forall v \in T_{\hat{x}}G, \tag{5.26}$$

where $d_1 E(\hat{x}, x)$ is the differential with respect to the first argument of E, and $d_1 E(\hat{x}, x) \circ v$ is the directional derivative of E along $v \in T_{\hat{x}}G$. Letting

$$\nabla_{\hat{x}} E(\hat{x}, x) = (\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x), \nabla_{\hat{m}} E(\hat{x}, x)), \tag{5.27}$$

where $\nabla_{\hat{m}} E(\hat{x}, x) = (\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^1} E(\hat{x}, x), \ldots, \boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^\ell} E(\hat{x}, x))$, from (5.23) the gradient of E with respect to $\hat{x}$ satisfies

$$\begin{aligned} \langle \nabla_{\hat{x}} E(\hat{x}, x), v \rangle_{\hat{x}} \;&=\; \langle \boldsymbol{\nabla}_{\hat{\mathbf{T}}_{ae}} E(\hat{x}, x) \mathbf{T}_{ae}^{-1}, \mathbf{V}\mathbf{T}_{ae}^{-1} \rangle \\ &\quad + \sum_{j=1}^{\ell} \mathbf{a}_j^{\mathsf{T}} \boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^j} E(\hat{x}, x), \end{aligned} \tag{5.28}$$

for all $v = (\mathbf{V}, a) \in T_{\hat{x}}G$.

It remains to find explicit expressions for $\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x)$ and $\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^j} E(\hat{x}, x)$. This is done in the following proposition.

Proposition 5.1. *Consider the error function, E, given in* (5.25) *along with the Riemannian metric defined in* (5.23). *Define $\tilde{\mathbf{y}}_j = \mathbf{T}_{ab}^{-1}\mathbf{m}_a^j - \mathbf{T}_{ae}^{-1}\hat{\mathbf{m}}_a^j \in \mathcal{M}_0$. Then,*

$$\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x) = \boldsymbol{\Lambda}\mathbf{T}_{ae}, \tag{5.29}$$

and

$$\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^j} E(\hat{x}, x) = -\boldsymbol{p}(\mathbf{T}_{ae}^{-\mathsf{T}}\tilde{\mathbf{y}}_j), \tag{5.30}$$

where $\boldsymbol{\Lambda} = \boldsymbol{\mathcal{P}}(\sum_{j=1}^{\ell} \mathbf{T}_{ae}^{-\mathsf{T}}\tilde{\mathbf{y}}_j \hat{\mathbf{m}}_a^{j\mathsf{T}})$.

Proof Define a curve $c : \mathbb{R} \to G$ such that $c(\epsilon) = (\boldsymbol{\Gamma}(\epsilon), \gamma(\epsilon))$, where $\boldsymbol{\Gamma}(\epsilon) \in SE(3)$ and $\gamma(\epsilon) \in \mathcal{M}_1^\ell$. Further define a tangent vector $v = (\mathbf{S}\mathbf{T}_{ae}, \zeta) \in T_{\hat{x}}G$ where $\mathbf{S} \in \mathfrak{se}(3)$

and $\zeta \in \mathcal{M}_0^\ell$. Let $\mathbf{\Gamma}(\epsilon) = \exp(\epsilon \mathbf{S})\mathbf{T}_{ae}$ and let $\gamma(\epsilon) = (\hat{\mathbf{m}}_a^1 + \epsilon\boldsymbol{\zeta}_1, \ldots, \hat{\mathbf{m}}_a^\ell + \epsilon\boldsymbol{\zeta}_\ell)$ such that $c(0) = \hat{x}$ and $\dot{c}(0) = v$. Then, the directional derivative of E with respect to $\hat{x}$ along v is given by [42]

$$d_1 E(\hat{x}, x) \circ v = \frac{\mathrm{d}}{\mathrm{d}\epsilon} E(c(\epsilon), x)\bigg|_{\epsilon=0}. \tag{5.31}$$

Substituting $c(\epsilon)$ into E it follows from (5.26) that

$$\begin{aligned}\langle \nabla_{\hat{x}} E(\hat{x}, x), v\rangle_x &= \mathrm{tr}(\textstyle\sum_{j=1}^{\ell} \hat{\mathbf{m}}_a^j \tilde{\mathbf{y}}_j^{\mathsf{T}} \mathbf{T}_{ae}^{-1} \mathbf{S}) - \sum_{j=1}^{\ell} \boldsymbol{\zeta}_j^{\mathsf{T}} \mathbf{T}_{ae}^{-\mathsf{T}} \tilde{\mathbf{y}}_j \\ &= \langle \boldsymbol{\mathcal{P}}(\mathbf{T}_{ae}^{-\mathsf{T}} \tilde{\mathbf{y}}_j \hat{\mathbf{m}}_a^{j\mathsf{T}}), \mathbf{S}\rangle - \sum_{j=1}^{\ell} \boldsymbol{\zeta}_j^{\mathsf{T}} \boldsymbol{p}(\mathbf{T}_{ae}^{-\mathsf{T}} \tilde{\mathbf{y}}_j).\end{aligned}$$

Comparing the above to (5.28), $\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x) = \mathbf{\Lambda}\mathbf{T}_{ae}$ and $\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^j} E(\hat{x}, x) = -\boldsymbol{p}(\mathbf{T}_{ae}^{-\mathsf{T}} \tilde{\mathbf{y}}_j)$, where $\mathbf{\Lambda} = \boldsymbol{\mathcal{P}}(\mathbf{T}_{ae}^{-\mathsf{T}} \tilde{\mathbf{y}}_j \hat{\mathbf{m}}_a^{j\mathsf{T}})$. $\square$

5.4.3 Observer Dynamics and Error Dynamics

Now that the SLAM group has been identified, a Riemannian metric defined, and a cost function determined, a gradient-based observer for SLAM can be constructed by the methods proposed in [14] and discussed in Chapter 3. The gradient-based observer for SLAM is abbreviated as GBO SLAM for the remainder of the chapter. By (3.19), GBO SLAM is given by

$$\dot{\hat{x}} = (\mathbf{T}_{ae}\mathbf{A}_b, 0) - K(\nabla_{\hat{x}} E(\hat{x}, x)), \tag{5.32}$$

where the symmetric positive definite gain mapping $K : T_{\hat{x}}G \to T_{\hat{x}}G$ has been added. The form of the observer shares the same structure as the SLAM algorithm proposed in [76], in that it is composed of a copy of the SLAM kinematics followed by an innovation term. The introduction of the gain mapping is inspired by [16], where the use of gain mappings for observer design on Lie groups is considered. The gain mapping $K(\cdot)$ allows for cross coupling between the innovations in the pose and map estimates. However, to simplify the analysis in later sections, coupling between pose and map innovations will not be considered, and thus $K(\nabla_{\hat{x}} E(\hat{x}, x))$ may be written as $K(\nabla_{\hat{x}} E(\hat{x}, x)) = (\boldsymbol{K}_T(\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x)), K_m(\nabla_m E(\hat{x}, x)))$, where $\boldsymbol{K}_T : T_{\mathbf{T}_{ae}} SE(3) \to T_{\mathbf{T}_{ae}} SE(3)$ and $K_m : \mathcal{M}_0^\ell \to \mathcal{M}_0^\ell$ are symmetric positive definite linear operators.

Consequently, (5.32) may be written in terms of $\mathbf{T}_{ae}$ and $\mathbf{m}_a^j$ as

$$\dot{\mathbf{T}}_{ae} = \mathbf{T}_{ae}\mathbf{A}_b - \boldsymbol{K}_T(\boldsymbol{\nabla}_{\mathbf{T}_{ae}} E(\hat{x}, x)), \tag{5.33a}$$

$$\dot{\hat{\mathbf{m}}}_a^j = -\textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^i} E(\hat{x}, x)), \tag{5.33b}$$

where $\boldsymbol{K}_m^{ji} : \mathcal{M}_0 \to \mathcal{M}_0$ is such that

$$K_m(\nabla_m E(\hat{x}, x)) = (\textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{1i}(\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^i} E(\hat{x}, x)), \ldots, \sum_{i=1}^{\ell} \boldsymbol{K}_m^{\ell i}(\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^i} E(\hat{x}, x))). \tag{5.34}$$

The gain mapping $K_m(\cdot)$ allows for cross coupling between landmark position innovations. In practice, this allows for corrections to the entire map after only a single landmark has been observed.

It was established in [14] that left invariant system dynamics coupled with a right invariant error function and a gradient defined with respect to a right invariant Riemannian metric yield autonomous error dynamics for the gradient-based observer. The error function, E, that has been chosen is not right invariant, which is to say that $E(x \star z, y \star z) \neq E(x, y)\ \forall x, y, z \in G$. Consequently, the stability results presented in [14] do not hold for the GBO SLAM algorithm considered in this chapter. However, because it is assumed that the landmarks remain stationary, that is $\dot{\mathbf{m}}_a^j = \mathbf{0}\ \forall j \in \{1, \ldots, \ell\}$, the error dynamics can be shown to be autonomous. For this purpose, define the group error as $\tilde{x} = \hat{x} \star x^{-1} = (\tilde{\mathbf{T}}, \tilde{m}) \in G$ such that $\tilde{\mathbf{T}} = \mathbf{T}_{ae}\mathbf{T}_{ab}^{-1}$ and $\tilde{m} = (\tilde{\mathbf{m}}_a^1, \ldots, \tilde{\mathbf{m}}_a^\ell)$, where $\underline{\tilde{\mathbf{m}}}_a^j = \underline{\hat{\mathbf{m}}}_a^j - \underline{\mathbf{m}}_a^j$. The autonomous nature of the dynamics associated with $\tilde{x}$ is established in the following proposition.

Proposition 5.2. *Let* $\mathbf{s}_j = (\tilde{\mathbf{T}} - \mathbf{1})\mathbf{m}_a^j - \boldsymbol{p}(\tilde{\mathbf{m}}_a^j) \in \mathcal{M}_0$. *Then, the error function may be written as* $E(\hat{x}, x) = \tilde{E}(\tilde{x})$, *where* $\tilde{E} : G \to \mathbb{R}^+$ *is given by*

$$\tilde{E}(\tilde{x}) = \tfrac{1}{2} \textstyle\sum_{j=1}^{\ell} ||(\tilde{\mathbf{T}} - \mathbf{1})\mathbf{m}_a^j - \boldsymbol{p}(\tilde{\mathbf{m}}_a^j)||_2^2 = \tfrac{1}{2} \sum_{j=1}^{\ell} \mathbf{s}_j^\mathsf{T} \mathbf{s}_j. \tag{5.35}$$

Moreover, assuming that $\boldsymbol{K}_T(\cdot)$ *is right invariant, that is* $\boldsymbol{K}_T(\mathbf{V})\mathbf{U} = \boldsymbol{K}_T(\mathbf{V}\mathbf{U})\ \forall \mathbf{V}, \mathbf{U} \in T_{\mathbf{T}_{ae}} SE(3)$, *the dynamics associated with* $\tilde{x}$ *are autonomous and are given by*

$$\dot{\tilde{\mathbf{T}}} = -\boldsymbol{K}_T(\boldsymbol{\nabla}_{\tilde{\mathbf{T}}} \tilde{E}(\tilde{x})) \tag{5.36a}$$

$$\dot{\tilde{\mathbf{m}}}_a^j = -\textstyle\sum_{j=1}^{\ell} \boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^i} \tilde{E}(\tilde{x})), \tag{5.36b}$$

where $\boldsymbol{\nabla}_{\tilde{\mathbf{T}}} \tilde{E}(\tilde{x}) = \boldsymbol{\Lambda}\tilde{\mathbf{T}}$ *and* $\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^j} \tilde{E}(\tilde{x}) = -\mathbf{s}_j$.

Proof Applying (5.10) to (5.25) gives

$$\begin{aligned} E(\hat{x}, x) &= \tfrac{1}{2}\textstyle\sum_{j=1}^{\ell} ||\mathbf{T}_{ab}^{-1}\mathbf{m}_a^j - \mathbf{T}_{ae}^{-1}\hat{\mathbf{m}}_a^j||_2^2 \\ &= \tfrac{1}{2}\textstyle\sum_{j=1}^{\ell} ||\mathbf{T}_{ae}^{-1}(\tilde{\mathbf{T}}\mathbf{m}_a^j - \hat{\mathbf{m}}_a^j)||_2^2 \\ &= \tfrac{1}{2}\textstyle\sum_{j=1}^{\ell} ||\tilde{\mathbf{T}}\mathbf{m}_a^j - \hat{\mathbf{m}}_a^j||_2^2. \end{aligned}$$

From the definition of $\tilde{\mathbf{m}}_a^j$, $\hat{\mathbf{m}}_a^j$ may be written as $\hat{\mathbf{m}}_a^j = \boldsymbol{p}(\tilde{\mathbf{m}}_a^j) + \mathbf{m}_a^j$. Thus, $E(\hat{x}, x) = \tilde{E}(\tilde{x}) = \frac{1}{2}\sum_{j=1}^{\ell} ||\mathbf{s}_j||_2^2$. By (5.3), (5.33a), and (5.29) the time derivative of $\tilde{\mathbf{T}}$ is

$$\begin{aligned} \dot{\tilde{\mathbf{T}}} &= \dot{\mathbf{T}}_{ae}\mathbf{T}_{ab}^{-1} + \mathbf{T}_{ae}\dot{\mathbf{T}}_{ab}^{-1} \\ &= \mathbf{T}_{ae}\mathbf{A}_b\mathbf{T}_{ab}^{-1} - \boldsymbol{K}_T(\boldsymbol{\Lambda}\mathbf{T}_{ae})\mathbf{T}_{ab}^{-1} - \mathbf{T}_{ae}\mathbf{A}_b\mathbf{T}_{ab}^{-1} \\ &= -\boldsymbol{K}_T(\boldsymbol{\Lambda}\mathbf{T}_{ae})\mathbf{T}_{ab}^{-1}. \end{aligned} \tag{5.37}$$

With the assumption that $\boldsymbol{K}_T(\cdot)$ is right invariant it follows that $\dot{\tilde{\mathbf{T}}} = -\boldsymbol{K}_T(\boldsymbol{\Lambda}\tilde{\mathbf{T}})$. Applying the fact that $\hat{\mathbf{m}}_a^j = \boldsymbol{p}(\tilde{\mathbf{m}}_a^j) + \mathbf{m}_a^j$, it follows that $\dot{\tilde{\mathbf{m}}}_a^j = \boldsymbol{p}(\dot{\tilde{\mathbf{m}}}_a^j) = \dot{\hat{\mathbf{m}}}_a^j$. Then from (5.33b) and (5.30), the time derivative of $\tilde{\mathbf{m}}_a^j$ satisfies

$$\begin{aligned} \dot{\tilde{\mathbf{m}}}_a^j &= -\textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^i} E(\hat{x}, x)) \\ &= \textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\boldsymbol{p}(\mathbf{T}_{ae}^{-\mathsf{T}}\tilde{\mathbf{y}}_i)). \end{aligned}$$

Applying the identity given in (5.11) it follows that

$$\begin{aligned} \dot{\tilde{\mathbf{m}}}_a^j &= \textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\tilde{\mathbf{T}}\mathbf{m}_a^i - \hat{\mathbf{m}}_a^i) && (5.38) \\ &= \textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}((\tilde{\mathbf{T}} - \mathbf{1})\mathbf{m}_a^i - \boldsymbol{p}(\tilde{\mathbf{m}}_a^i)) && (5.39) \\ &= \textstyle\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\mathbf{s}_i). && (5.40) \end{aligned}$$

To show that the error dynamics are autonomous, it is sufficient to show that $\boldsymbol{\Lambda}$ may be written in terms of $\tilde{\mathbf{T}}$, $\tilde{\mathbf{m}}_a^j$, and $\mathbf{m}_a^j$. Applying (5.8) and the fact that $\hat{\mathbf{m}}_a^j = \boldsymbol{p}(\tilde{\mathbf{m}}_a^j) + \mathbf{m}_a^j$ gives

$$\begin{aligned} \boldsymbol{\Lambda} &= \boldsymbol{\mathcal{P}}(\textstyle\sum_{j=1}^{\ell} \mathbf{T}_{ae}^{-\mathsf{T}}(\mathbf{T}_{ab}^{-1}\mathbf{m}_a^j - \mathbf{T}_{ae}^{-1}\hat{\mathbf{m}}_a^j)\hat{\mathbf{m}}_a^{j\mathsf{T}}) \\ &= \boldsymbol{\mathcal{P}}(\textstyle\sum_{j=1}^{\ell}(\tilde{\mathbf{T}}\mathbf{m}_a^j - \hat{\mathbf{m}}_a^j)\hat{\mathbf{m}}_a^{j\mathsf{T}}) \\ &= \boldsymbol{\mathcal{P}}(\textstyle\sum_{j=1}^{\ell} \mathbf{s}_j(\mathbf{m}_a^j + \boldsymbol{p}(\tilde{\mathbf{m}}_a^j))^{\mathsf{T}}), \end{aligned} \tag{5.41}$$

and therefore the error dynamics are autonomous.

Following the same method used in the proof of Proposition 5.1, it can be shown

that $\boldsymbol{\nabla}_{\tilde{\mathbf{T}}}\tilde{E}(\tilde{x}) = \boldsymbol{\Lambda}\tilde{\mathbf{T}}$ and $\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^j}\tilde{E}(\tilde{x}) = -\mathbf{s}_j$. Thus, $\dot{\tilde{\mathbf{T}}} = -\boldsymbol{K}_T(\boldsymbol{\nabla}_{\tilde{\mathbf{T}}}\tilde{E}(\tilde{x}))$ and $\dot{\tilde{\mathbf{m}}}_a^j = -\sum_{j=1}^{\ell}\boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^i}\tilde{E}(\tilde{x}))$, which completes the proof. □

5.5 Bias Estimation

In the previous section it was assumed that $\mathbf{A}_b$ was available for use in the observer equations. Now, consider the case when $\mathbf{A}_b$ is measured as $\mathbf{A}_b^y = \mathbf{A}_b + \mathbf{B} \in \mathfrak{se}(3)$, where $\mathbf{B} \in \mathfrak{se}(3)$ is a constant unknown bias such that

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_\omega^\times & \mathbf{b}_v \\ \mathbf{0} & 0 \end{bmatrix}, \tag{5.42}$$

and $\mathbf{b}_\omega, \mathbf{b}_v \in \mathbb{R}^3$. In these instances it is necessary to modify the observer equations of (5.33) by including a bias estimator to estimate the unknown bias $\mathbf{B}$. This method and has been successfully employed in the context of attitude estimation [10], pose estimation [27], on general Lie groups [16], and was previously discussed in Chapter 4.

Let $\hat{\mathbf{B}} \in \mathfrak{se}(3)$ denote the estimate of the bias $\mathbf{B}$. Consider an observer of the form

$$\begin{aligned}
\dot{\mathbf{T}}_{ae} &= \mathbf{T}_{ae}(\mathbf{A}_b^y - \hat{\mathbf{B}}) - \boldsymbol{K}_T(\boldsymbol{\nabla}_{\mathbf{T}_{ae}}E(\hat{x}, x)), && \text{(5.43a)}\\
\dot{\hat{\mathbf{m}}}_a^j &= -\textstyle\sum_{i=1}^{\ell}\boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\hat{\mathbf{m}}_a^j}E(\hat{x}, x)), && \text{(5.43b)}\\
\dot{\hat{\mathbf{B}}} &= \rho\boldsymbol{\mathcal{P}}(\mathbf{T}_{ae}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{T}_{ae}^{-\mathsf{T}}), && \text{(5.43c)}
\end{aligned}$$

where $\rho \in (0, \infty)$. Define the error in the bias estimate as $\tilde{\mathbf{B}} = \mathbf{B} - \hat{\mathbf{B}} \in \mathfrak{se}(3)$. Following a procedure similar to Proposition 5.2, the error dynamics can be found to be nonautonomous and of the form

$$\begin{aligned}
\dot{\tilde{\mathbf{T}}} &= \mathbf{T}_{ae}\tilde{\mathbf{B}}\mathbf{T}_{ae}^{-1}\tilde{\mathbf{T}} - \boldsymbol{K}_T(\boldsymbol{\nabla}_{\tilde{\mathbf{T}}}\tilde{E}(\tilde{x})), && \text{(5.44a)}\\
\dot{\tilde{\mathbf{m}}}_a^j &= -\textstyle\sum_{i=1}^{\ell}\boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^j}\tilde{E}(\tilde{x})), && \text{(5.44b)}\\
\dot{\tilde{\mathbf{B}}} &= -\rho\boldsymbol{\mathcal{P}}(\mathbf{T}_{ae}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{T}_{ae}^{-\mathsf{T}}). && \text{(5.44c)}
\end{aligned}$$

5.6 Stability of GBO SLAM

The convergence of the observers presented in (5.33) and (5.43) will now be assessed. Theorem 5.3 deals with the convergence properties of (5.33), while Theorem

5.4 deals with the convergence of (5.43).

Theorem 5.3. *All trajectories of $\tilde{x}$ satisfying* (5.36) *asymptotically approach the set $M = \{\tilde{x} \in G \mid \tilde{E}(\tilde{x}) = 0\}$. Moreover, there exists constant $\mathbf{T} \in SE(3)$ such that, in M, $\mathbf{T}_{ae} = \mathbf{T}\mathbf{T}_{ab}$ and $\hat{\mathbf{m}}_a^j = \mathbf{T}\mathbf{m}_a^j$.*

Proof Consider the lower bounded function

$$V_1(\tilde{x}) = \tilde{E}(\tilde{x}). \tag{5.45}$$

Taking the derivative with respect to time of V_1 gives

$$\begin{aligned}
\dot{V}_1 &= \langle \nabla_{\tilde{x}} E(\tilde{x}), \dot{\tilde{x}} \rangle_{\tilde{x}} \\
&= -\langle \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}), \boldsymbol{K}_T(\boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x})) \rangle_{\tilde{\mathbf{T}}} \\
&\quad - \langle \nabla_{\tilde{m}} \tilde{E}(\tilde{x}), K_m(\nabla_{\tilde{m}} \tilde{E}(\tilde{x})) \rangle_{\tilde{m}} \\
&\leq -\underline{k}_T \langle \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}), \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}) \rangle_{\tilde{\mathbf{T}}} \\
&\quad - \underline{k}_m \langle \nabla_{\tilde{m}} \tilde{E}(\tilde{x}), \nabla_{\tilde{m}} \tilde{E}(\tilde{x}) \rangle_{\tilde{m}} \\
&= -\underline{k}_T \langle \boldsymbol{\Lambda}, \boldsymbol{\Lambda} \rangle - \underline{k}_m \textstyle\sum_{j=1}^{\ell} \mathbf{s}_j^{\mathsf{T}} \mathbf{s}_j,
\end{aligned}$$

where $\underline{k}_T$ and $\underline{k}_m$ are the minimum eigenvalues of $\boldsymbol{K}_T$ and K_m, respectively. Consequently, $V_1(t) \leq V_1(0)$ and the set $\Omega = \{\tilde{x} \in G \mid V_1(t) \leq V_1(0)\}$ is compact and positively invariant. Let $\mathcal{E} = \{\tilde{x} \in \Omega \mid \dot{V}_1 = 0\}$. As $\dot{V}_1$ is bounded from below by $-\bar{k}_T \langle \boldsymbol{\Lambda}, \boldsymbol{\Lambda} \rangle_{\mathfrak{se}(3)} - \bar{k}_m \sum_{j=1}^{\ell} \mathbf{s}_j^{\mathsf{T}} \mathbf{s}_j$, where $\bar{k}_T$ and $\bar{k}_m$ denote the maximum eigenvalues of $\boldsymbol{K}_T$ and K_m, it follows that $\mathcal{E} = \{\tilde{x} \in \Omega \mid \mathbf{s}_j = \mathbf{0}, \boldsymbol{\Lambda} = \mathbf{0}\}$. However, $\mathbf{s}_j = \mathbf{0}$ implies that $\boldsymbol{\Lambda} = \mathbf{0}$ and, via (5.35), $V_1 = 0$. Therefore, $\mathcal{E} = \{\tilde{x} \in \Omega \mid V_1 = 0\}$. As $\mathcal{E}$ itself is an invariant set, the largest invariant set in $\mathcal{E}$ is the set $M = \mathcal{E}$. By LaSalle's invariance principle all trajectories of $\tilde{x}$ starting in Ω converge to the set $M = \{\tilde{x} \in \Omega \mid V_1 = 0\}$ as $t \longrightarrow \infty$.

Trajectories in M satisfy $V_1 \equiv 0$ and thus $\mathbf{s}_j \equiv \mathbf{0}$. Therefore, $\dot{\tilde{\mathbf{m}}}_a^j \equiv \mathbf{0}$ and $\dot{\tilde{\mathbf{T}}} \equiv \mathbf{0}$, which implies that $\tilde{\mathbf{m}}_a^j$ and $\tilde{\mathbf{T}}$ are constant in M. Moreover, $\mathbf{s}_j \equiv \mathbf{0}$ implies that $\hat{\mathbf{m}}_a^j \equiv \tilde{\mathbf{T}}\mathbf{m}_a^j$. Let $\mathbf{T} \in SE(3)$ denote the constant value that $\tilde{\mathbf{T}}$ reaches in M. Then, the limit points of trajectories of $\tilde{x}$ satisfy $\mathbf{T}_{ae} = \mathbf{T}\mathbf{T}_{ab}$ and $\hat{\mathbf{m}}_a^j = \mathbf{T}\mathbf{m}_a^j$. Thus the limit points of trajectories of $\tilde{x}$ satisfy $\mathbf{T}_{ae} = \mathbf{T}\mathbf{T}_{ab}$ and $\hat{\mathbf{m}}_a^j = \mathbf{T}\mathbf{m}_a^j$. In other words, as $t \to \infty$ the map and pose estimates approach their true counterparts transformed by some constant pose transformation. □

Note that the results of Theorem 5.3 do not depend on the number of landmarks. In addition, no restrictions have yet been placed on the linear independence of the

landmark positions, unlike in [10, 26] where $SO(3)$ and $SE(3)$ observers are considered. For example, in [26] it is assumed that there exists three or more landmarks whose positions are not collinear. This guarantees the existence of four isolated equilibria and the local asymptotic stability of the desired equilibrium point. This can be interpreted as an observability condition, as is done in [27]. Moreover, a variation of this condition was discussed in Chapter 4 where it was manifested as a rank condition on the matrix $\mathbf{M}_a$ in (4.14). However, the estimation problem currently considered is inherently unobservable, as discussed earlier in Section 5.2. Further, there exists no isolated equilibria or even a desired equilibrium point, but rather a desired invariant set M. Consequently, no landmark restrictions are necessary for the results of Theorem 1. This is not to say, however, that the number and properties of the landmarks have no effect in practice. In reality, nonideal input measurements will cause drift and other undesirable effects in the state estimates. If $\ell < 3$ then drift in certain directions of the state-space will not cause an increase in the estimation error E and consequently the innovation terms will not correct the drift. It may be possible to strengthen the results of Theorem 5.3 by placing restrictions on the number and properties of the landmarks. However, this will be left as a possible extension to the current work. Instead, it will now be shown that undesirable effects associated with constant bias in the input measurements can be countered provided the landmarks satisfy a certain linear independence condition.

Theorem 5.4. *Consider trajectories of* $(\tilde{x}, \tilde{\mathbf{B}})$ *under* (5.44) *and assume the following:*

(A1) Trajectories of $\tilde{x}(t)$ *remain bounded for all* $t \geq 0$.

(A2) Trajectories of $\mathbf{T}_{ab}(t)$ *and* $\mathbf{A}_b(t)$ *remain bounded for all* $t \geq 0$.

(A3) The number of total landmarks is greater than or equal to three and

$$span\{\mathbf{m}_a^1, \ldots, \mathbf{m}_a^\ell\} = \mathcal{M}_1.$$

Then, trajectories of $(\tilde{x}, \tilde{\mathbf{B}})$ *asymptotically approach the set* $\bar{M} = \{(\tilde{x}, \tilde{\mathbf{B}}) \in G \times \mathfrak{se}(3) \mid \tilde{E}(\tilde{x}) = 0,\ \tilde{\mathbf{B}} = \mathbf{0}\}$. *Moreover, there exists constant* $\mathbf{T} \in SE(3)$ *such that, in* $\bar{M}$, $\mathbf{T}_{ae} = \mathbf{T}\mathbf{T}_{ab}$ *and* $\hat{\mathbf{m}}_a^j = \mathbf{T}\mathbf{m}_a^j$.

Proof Consider the lower bounded function

$$V_2(\tilde{x}, \tilde{\mathbf{B}}) = \tilde{E}(\tilde{x}) + \frac{1}{2\rho}\langle \tilde{\mathbf{B}}, \tilde{\mathbf{B}} \rangle, \tag{5.46}$$

where $\tilde{E}$ is given in (5.35). Taking the derivative with respect to time of V_2 gives

$$\begin{aligned}
\dot{V}_2 &= \langle \nabla_{\tilde{x}} E(\tilde{x}), \dot{\tilde{x}} \rangle_{\tilde{x}} + 1/\rho \langle \tilde{\mathbf{B}}, \dot{\tilde{\mathbf{B}}} \rangle \\
&\leq -\underline{k}_T \langle \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}), \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}) \rangle_{\tilde{\mathbf{T}}} + \langle \boldsymbol{\Lambda}, \mathbf{T}_{ae} \tilde{\mathbf{B}} \mathbf{T}_{ae}^{-1} \rangle \\
&\quad - \underline{k}_m \langle \nabla_{\tilde{m}} \tilde{E}(\tilde{x}), \nabla_{\tilde{m}} \tilde{E}(\tilde{x}) \rangle_{\tilde{m}} - \langle \tilde{\mathbf{B}}, \mathbf{T}_{ae}^{\mathsf{T}} \boldsymbol{\Lambda} \mathbf{T}_{ae}^{-\mathsf{T}} \rangle \\
&= -\underline{k}_T \langle \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}), \boldsymbol{\nabla}_{\tilde{\mathbf{T}}} E(\tilde{x}) \rangle_{\tilde{\mathbf{T}}} \\
&\quad - \underline{k}_m \langle \nabla_{\tilde{m}} \tilde{E}(\tilde{x}), \nabla_{\tilde{m}} \tilde{E}(\tilde{x}) \rangle_{\tilde{m}} \\
&= -\underline{k}_T \langle \boldsymbol{\Lambda}, \boldsymbol{\Lambda} \rangle - \underline{k}_m \textstyle\sum_{j=1}^{\ell} \mathbf{s}_j^{\mathsf{T}} \mathbf{s}_j.
\end{aligned}$$

Therefore, $\dot{V}_2 \leq -\underline{k}_m \sum_{j=1}^{\ell} \mathbf{s}_j^{\mathsf{T}} \mathbf{s}_j = -2\underline{k}_m \tilde{E}(\tilde{x})$. It follows then that $V_2 \in \mathcal{L}_\infty$ and $\lim_{t\to\infty} V_2(t)$ exists and is finite. Thus, $\tilde{\mathbf{B}}$ remains bounded for all $t \geq 0$ and $\tilde{\mathbf{B}}$ converges to some constant value. In addition, it follows that $\tilde{E} \in \mathcal{L}_\infty$ and that $\lim_{t\to\infty} \tilde{E}$ exists and is finite [80, p. 181, Fact 1]. Moreover, $\int_0^\infty \tilde{E}(\tau)\mathrm{d}\tau \leq 1/(2\underline{k}_m)(V_2(0) - \lim_{t\to\infty} V_2(t)) < \infty$. As $\tilde{E} \geq 0$, it follows that $\tilde{E} \in \mathcal{L}_1$. The fact that $\tilde{E}$ converges along with $\tilde{E} \in \mathcal{L}_1$ implies that $\tilde{E} \to 0$ as $t \to \infty$ [80, pp. 181–182, Fact 3 and Fact 5].

The fact that $\tilde{E}(\tilde{x})$ approaches zero implies that $\lim_{t\to\infty} \mathbf{s}_j = \mathbf{0}$ for all $j \in \{1, \ldots, \ell\}$. By assumption (A1) and (5.41), $\lim_{t\to\infty} \mathbf{s}_j = \mathbf{0}$ implies that $\boldsymbol{\Lambda} \to \mathbf{0}$ as $t \to \infty$. With assumptions (A1) and (A2), it can be shown that $\ddot{\mathbf{s}}_j$ remains bounded and consequently $\dot{\mathbf{s}}_j$ is uniformly continuous. The application of Barbalat's Lemma yields $\dot{\mathbf{s}}_j \to \mathbf{0}$ as $t \to \infty$. Given that $\dot{\mathbf{s}}_j = \tilde{\mathbf{T}}\mathbf{T}_{ab}\tilde{\mathbf{B}}\mathbf{T}_{ab}^{-1}\mathbf{m}_a^j - \boldsymbol{K}_T(\boldsymbol{\Lambda}\tilde{\mathbf{T}})\mathbf{m}_a^j - \sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\mathbf{s}_i)$, assumption (A1), and the facts that $\boldsymbol{\Lambda} \to \mathbf{0}$ and $\mathbf{s}_j \to \mathbf{0}$, it follows that $\lim_{t\to\infty} \dot{\mathbf{s}}_j = \lim_{t\to\infty} \tilde{\mathbf{T}}\mathbf{T}_{ab}\tilde{\mathbf{B}}\mathbf{T}_{ab}^{-1}\mathbf{m}_a^j = \mathbf{0}$. Consequently, $\lim_{t\to\infty} \tilde{\mathbf{B}}\mathbf{y}_j = \mathbf{0}$ for all $j \in \{1, \ldots, \ell\}$, where $\mathbf{y}_j \triangleq \mathbf{T}_{ab}^{-1}\mathbf{m}_a^j \in \mathcal{M}_1$. It therefore follows that in the limit as $t \to \infty$ the equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ must be satisfied, where $\mathbf{x} = [\ \tilde{\mathbf{b}}_\omega^{\mathsf{T}}\ \tilde{\mathbf{b}}_v^{\mathsf{T}}\ ]^{\mathsf{T}}$ and

$$\mathbf{A} = \begin{bmatrix} -\underline{\mathbf{y}}_1^{\times} & \mathbf{1} \\ \vdots & \vdots \\ -\underline{\mathbf{y}}_\ell^{\times} & \mathbf{1} \end{bmatrix}. \tag{5.47}$$

By assumption (A3), the set of landmark positions $\mathbf{m}_a^j$ spans $\mathcal{M}_1$, which implies that $\mathrm{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_\ell\} = \mathcal{M}_1$ and $\mathrm{span}\{\underline{\mathbf{y}}_1, \ldots, \underline{\mathbf{y}}_\ell\} = \mathbb{R}^3$. The matrix $\mathbf{A}$ has full column rank provided $\mathrm{span}\{\underline{\mathbf{y}}_1, \ldots, \underline{\mathbf{y}}_\ell\} = \mathbb{R}^3$. Thus, $\mathbf{A}$ has full column rank and $\mathbf{A}\mathbf{x} = \mathbf{0}$ implies that $\mathbf{x} = \mathbf{0}$, which in turn, implies that $\tilde{\mathbf{B}} = \mathbf{0}$. This shows that $\tilde{\mathbf{B}} \to \mathbf{0}$ as $t \to \infty$.

Suppose $(\bar{x}(t), \bar{\mathbf{B}}(t))$ is a solution of (5.44). Then $\bar{x}(t) = (\bar{\mathbf{T}}(t), \bar{m}(t))$ is also a

solution of

$$\dot{\tilde{\mathbf{T}}} = \mathbf{T}_{ae}\bar{\mathbf{B}}(t)\mathbf{T}_{ae}^{-1}\tilde{\mathbf{T}} - \boldsymbol{K}_T(\boldsymbol{\nabla}_{\tilde{\mathbf{T}}}\tilde{E}(\tilde{x})), \tag{5.48a}$$

$$\dot{\tilde{\mathbf{m}}}_a^j = -\sum_{i=1}^{\ell} \boldsymbol{K}_m^{ji}(\boldsymbol{\nabla}_{\tilde{\mathbf{m}}_a^j}\tilde{E}(\tilde{x})), \tag{5.48b}$$

As $\mathbf{T}_{ae}$ and $\tilde{\mathbf{T}}$ are bounded by assumption and $\bar{\mathbf{B}}(t) \to \mathbf{0}$ it follows that (5.48) is asymptotically autonomous to (5.36) and thus the positive limit sets of solutions of (5.48) are invariant under (5.36). Consequently, the positive limit sets of (5.44) are invariant under (5.36) [81, Theorem 4]. The results of Theorem 5.3 indicate that positive limit sets of (5.36) satisfy $\mathbf{T}_{ae} = \mathbf{T}\mathbf{T}_{ab}$ and $\hat{\mathbf{m}}_a^j = \mathbf{T}\mathbf{m}_a^j$ for some constant $\mathbf{T} \in SE(3)$. Thus, positive limit sets of (5.44) satisfy this same condition. This completes the proof. □

5.7 Numerical Example

In this section, the GBO SLAM algorithm is implemented in simulation. The simulation parameters in this section are inspired by the simulation conducted in [76]. The goal of this section is to verify the convergence properties of GBO SLAM, both with and without bias estimation. Two cases are considered. In the first case, no noise or bias are added to the velocity and landmark position measurements. In the second case, bias is added to the velocity measurements. It is assumed that there exists $\ell = 3$ landmarks surrounding the vehicle. The landmark positions are given by

$$\underline{\mathbf{m}}_a^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ (m)}, \ \underline{\mathbf{m}}_a^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ (m)}, \text{ and, } \underline{\mathbf{m}}_a^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ (m)}, \tag{5.49}$$

which satisfy assumption (A3) in Theorem 5.4. The initial vehicle position is given by $\mathbf{r}_a^{ba}(0) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ and the initial vehicle attitude is $\mathbf{C}_{ab}(0) = \mathbf{1}$. The trajectory of the vehicle is determined by (5.3) where the angular and translational velocities are taken as $\boldsymbol{\omega}_b^{ba} = \mathbf{0}$ (rad/s) and

$$\mathbf{v}_b^{ba} = \begin{bmatrix} 0.05\cos(\pi/25t) \\ 0.05\sin(\pi/25t) \\ 0.05 \end{bmatrix} \text{ (m/s)}. \tag{5.50}$$

The pose estimate is initialized as $\mathbf{T}_{ae}(0) = \mathbf{T}_{ba}(0)$ and the map estimate is initialized with

$$\underline{\hat{\mathbf{m}}}_a^1 = \begin{bmatrix} 0.8 \\ 0.3 \\ 0.0 \end{bmatrix} \text{ (m)}, \ \underline{\hat{\mathbf{m}}}_a^2 = \begin{bmatrix} 0.4 \\ 1.3 \\ 0.0 \end{bmatrix} \text{ (m)}, \text{ and, } \underline{\hat{\mathbf{m}}}_a^3 = \begin{bmatrix} 0.0 \\ 0.2 \\ 0.8 \end{bmatrix} \text{ (m)}. \tag{5.51}$$

In this example, the symmetric positive definite gain K in (5.32) is taken to be the identity operator.

5.7.1 Case 1: No Bias

In the first simulation the GBO SLAM algorithm described by (5.32) is implemented. Recall, by Theorem 5.3, it is expected that the pose and map estimates will approach their true counterparts transformed by some constant pose transformation. The results of the simulation are shown in Figs. 5.1 and 5.2. The attitude and position error taken from

$$\tilde{\mathbf{T}} = \begin{bmatrix} \tilde{\mathbf{C}} & \tilde{\mathbf{r}} \\ \mathbf{0} & 1 \end{bmatrix}, \tag{5.52}$$

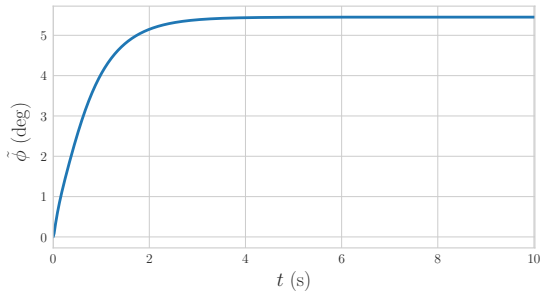
is shown in Fig. 5.1a and Fig. 5.1b, respectively. The attitude error $\tilde{\phi}$ is extracted from $\tilde{\mathbf{C}}$ by $\tilde{\phi} = \arccos(\frac{1}{2}(\text{tr}(\tilde{\mathbf{C}}) - 1))$. As predicted, the error in pose approaches a constant value. The cost function approaches zero as shown in Fig. 5.1c confirming the results of Theorem 5.3. The constant pose transformation between the estimated and true state can be seen visually in Fig. 5.2 where a 3D visualization of the simulation is shown.

5.7.2 Case 2: Bias Estimation

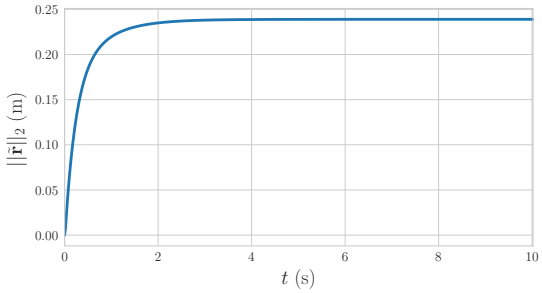
The GBO SLAM algorithm with bias estimation, given by (5.43), is now implemented. The goal is to verify the results of Theorem 5.4. Bias is introduced to the velocity measurement as described in (5.42) where $\mathbf{b}_\omega = \begin{bmatrix} 0.01 & -0.01 & 0.01 \end{bmatrix}^\mathsf{T}$ (rad/s) and $\mathbf{b}_v = \begin{bmatrix} 0.005 & 0.003 & -0.002 \end{bmatrix}^\mathsf{T}$ (m/s). The observer gain ρ is taken as $\rho = 1.0$. As in Sec. 4.5, (5.43c) is only implemented provided

$$\sqrt{\langle \nabla_{\hat{x}} E(\hat{x}, x), \nabla_{\hat{x}} E(\hat{x}, x) \rangle_{\hat{x}}} \leq 0.1, \tag{5.53}$$

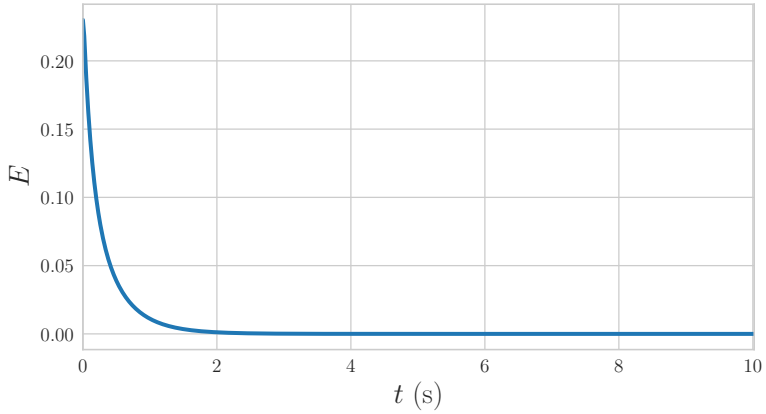
otherwise $\dot{\hat{\mathbf{B}}} = \mathbf{0}$. This is necessary to avoid integrator windup.



(a) Attitude error.



(b) Position error.



(c) Cost function, E.

Figure 5.1: GBO SLAM simulation results. Pose estimation error approaches a constant while the cost function approaches zero.

The results of the simulation are shown in Figs. 5.3, 5.4, and 5.5. As in the previous case, the errors in attitude in pose approach constants, while the cost function approaches zero. The errors associated with the bias estimate are plotted in Fig. 5.4, where it can be seen that the errors approach zero indicating successful bias estimation, a result predicted in Theorem 5.4. A 3D visualization of the simulation is shown in Fig. 5.5.

5.8 Experimental Results

GBO SLAM is now tested using the experimental data of [82]. The experimental setup consists of 20 reflective markers placed on the ground as well as a sensor head. The sensor head is equipped with an inertial measurement unit that provides measurements of angular velocity and translational acceleration while an onboard stereo camera observes the reflective markers. The translational acceleration measurements

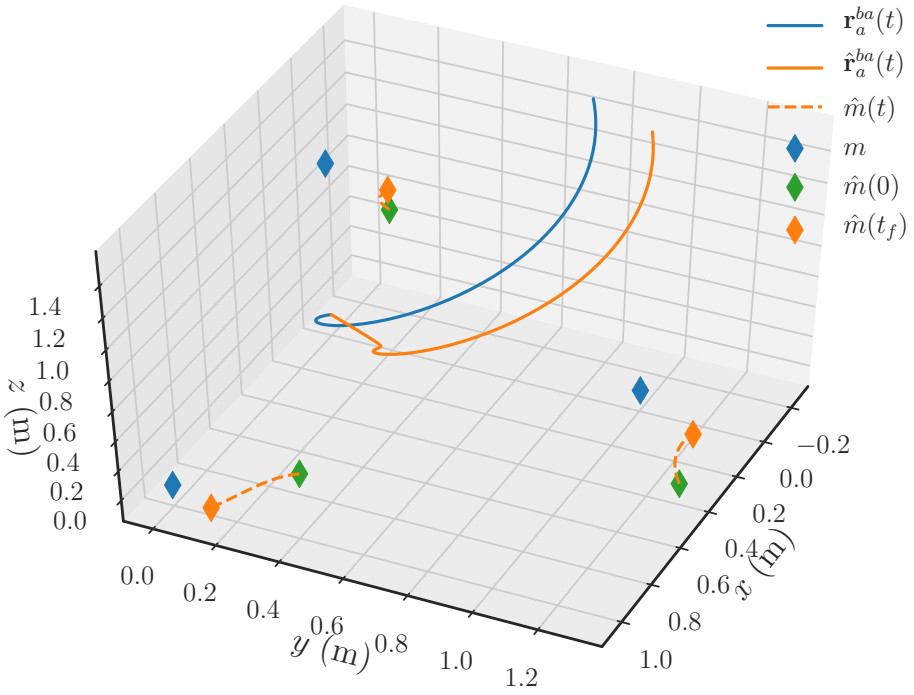
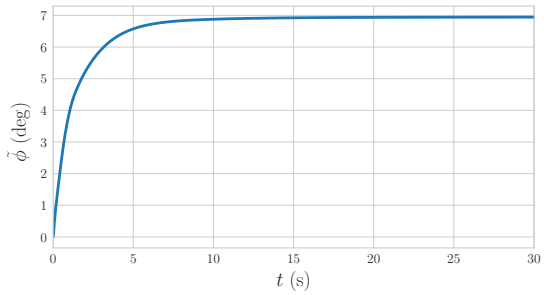


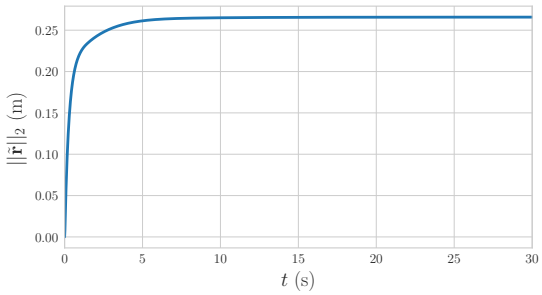
Figure 5.2: 3D visualization of GBO SLAM. The true vehicle position, shown in blue, as well as the estimated position, shown in orange, are shown. In addition, the map estimate, $\hat{m}$, at time $t = 0$ (s) and $t = t_f = 30$ (s) are also shown. Due to initialization errors the estimated map and pose are shifted by a constant transformation, mostly in the positive y direction.

are integrated thus generating artificial translational velocity measurements. Stereo features are associated with landmarks using ground truth data. Ground truth measurements of the landmark positions and pose of the sensor head are taken using a motion capture system. The SLAM algorithm is implemented offline after data has been acquired by the sensors. To emphasize the bias estimation capability of the proposed method large artificial biases are added to the angular and translational velocity after the true biases are negated using ground-truth data. The added artificial bias takes the form of (5.42), where $\mathbf{b}_\omega = [\ 0.016\ \ -0.01\ 0.014\ ]^\mathsf{T}$ (rad/s) and $\mathbf{b}_v = [\ 0.04\ 0.02\ \ -0.02\ ]^\mathsf{T}$ (m/s).

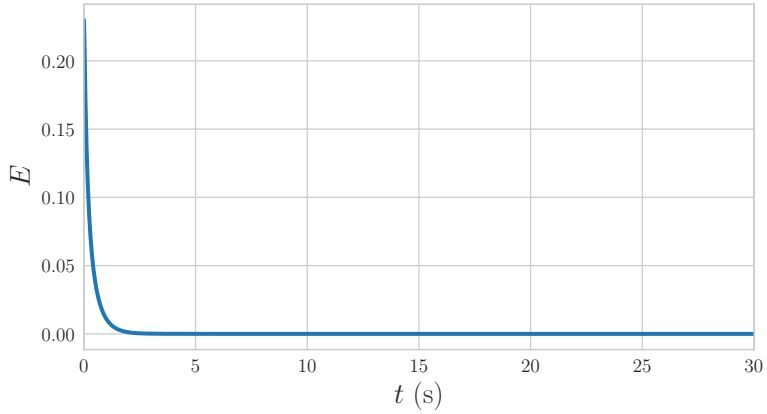
GBO SLAM is implemented in discrete time. It is possible to design Lie group observers directly in discrete time as is done in [83] and [84]. However, in this section the discrete-time system equations are found by directly discretizing the continuous-time system as is done in [10]. Let T denote the sample time and let $\mathbf{T}_{ae,k}$, $\hat{\mathbf{m}}_{a,k}^j$, and $\hat{\mathbf{B}}_k$ denote the state estimates at time $t = kT$, where $k \in \mathbb{N}$. Further, let $\mathbf{y}_{j,k} \in \mathcal{M}_1$



(a) Attitude error.

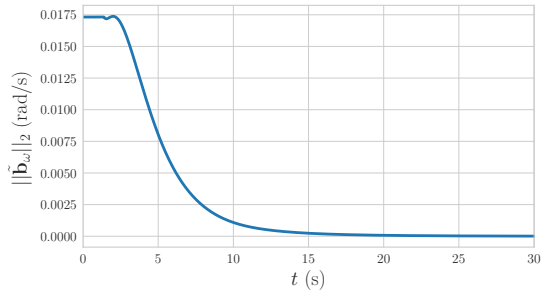


(b) Position error.

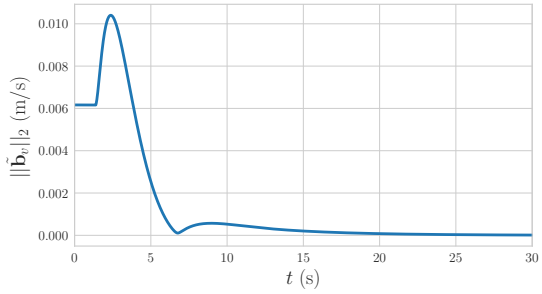


(c) Cost function, E.

Figure 5.3: GBO SLAM simulation results.



(a) Angular velocity bias estimation error.



(b) Translational velocity bias estimation error.

Figure 5.4: Error in the bias estimate.

be the measurement of the position of landmark j at time $t = kT$. The measurement $\mathbf{A}_b^y$ at time $t = kT$ is denoted $\mathbf{A}_{b,k}^y$. To implement (5.43) in discrete time the state

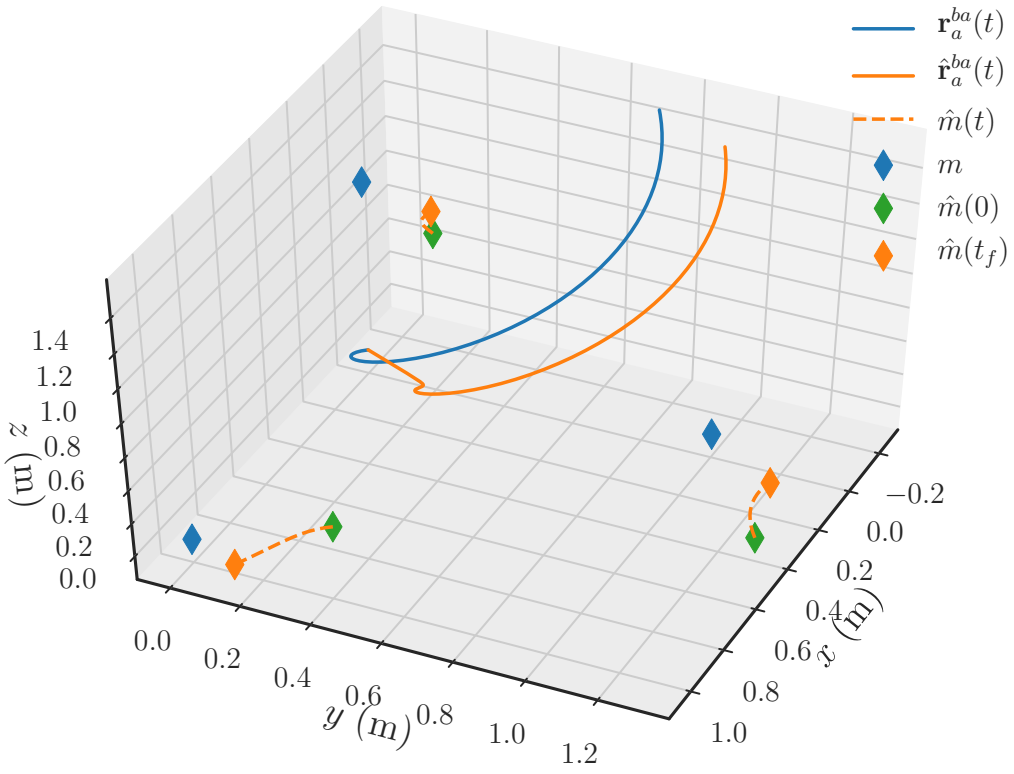


Figure 5.5: 3D visualization of GBO SLAM with bias estimation. Bias estimation is successful and the results mirror the results shown in Fig. 5.2. The true vehicle position is shown in blue, while the estimated position is shown in orange.

estimates are propagated forward by

$$\begin{aligned}
\mathbf{T}_{ae,k}^{-} &= \mathbf{T}_{ae,k-1}\exp(T(\mathbf{A}_{b,k-1}^{y} - \hat{\mathbf{B}}_{k-1})),\\
\hat{\mathbf{m}}_{a,k}^{j-} &= \hat{\mathbf{m}}_{a,k-1}^{j},\\
\hat{\mathbf{B}}_{k}^{-} &= \hat{\mathbf{B}}_{k-1},
\end{aligned}$$

and corrected by

$$\begin{aligned}
\mathbf{T}_{ae,k} &= \mathbf{T}_{ae,k}^{-}\exp(T\boldsymbol{K}_T((\mathbf{T}_{ae,k}^{-})^{-1}\boldsymbol{\Lambda}_k\mathbf{T}_{ae,k}^{-})),\\
\hat{\mathbf{m}}_{a,k}^{j} &= \hat{\mathbf{m}}_{a,k}^{j-} + T\textstyle\sum_{i\in\mathcal{O}_k}\boldsymbol{K}_m^{ji}(\boldsymbol{p}((\mathbf{T}_{ae,k}^{-})^{-\mathsf{T}}\tilde{\mathbf{y}}_{i,k})),\\
\hat{\mathbf{B}}_{k} &= \hat{\mathbf{B}}_{k}^{-} + \rho T\boldsymbol{\mathcal{P}}((\mathbf{T}_{ae,k}^{-})^{\mathsf{T}}\boldsymbol{\Lambda}_k(\mathbf{T}_{ae,k}^{-})^{-\mathsf{T}}),
\end{aligned}$$

where $\boldsymbol{\Lambda}_k = \boldsymbol{\mathcal{P}}(\sum_{j\in\mathcal{O}_k}(\mathbf{T}_{ae,k}^{-})^{-\mathsf{T}}\tilde{\mathbf{y}}_{j,k}\hat{\mathbf{m}}_{a,k}^{j\mathsf{T}})$, and $\mathcal{O}_k$ is the set of all landmarks that have been observed at time $t = Tk$. Taking the standard basis of $\mathfrak{se}(3)$, denoted $\{\mathbf{B}\}$, and

right translating by $\mathbf{T}_{ae}$ giving $\{\mathbf{BT}_{ae}\}$ yields a basis for $T_{\mathbf{T}_{ae}}SE(3)$ [16]. The gain $\boldsymbol{K}_T$ in this basis is taken to be $[K_T]_{\{\mathbf{BT}_{ae}\}} = \text{diag}\{1, 1, 1, 0.5, 0.5, 0.5\}$. The gain K_m in the standard basis of $\mathcal{M}_0^\ell$ is taken as $0.1\mathbf{1}$, while the value of ρ is chosen as $\rho = 0.03$. The estimates are initialized as $\mathbf{T}_{ae,0} = \mathbf{1}$ and $\hat{\mathbf{B}}_0 = \mathbf{0}$. Landmark estimates are initialized when first observed as $\hat{\mathbf{m}}_{a,k}^{j-} = \mathbf{T}_{ae,k}^{-}\mathbf{y}_{j,k}$.

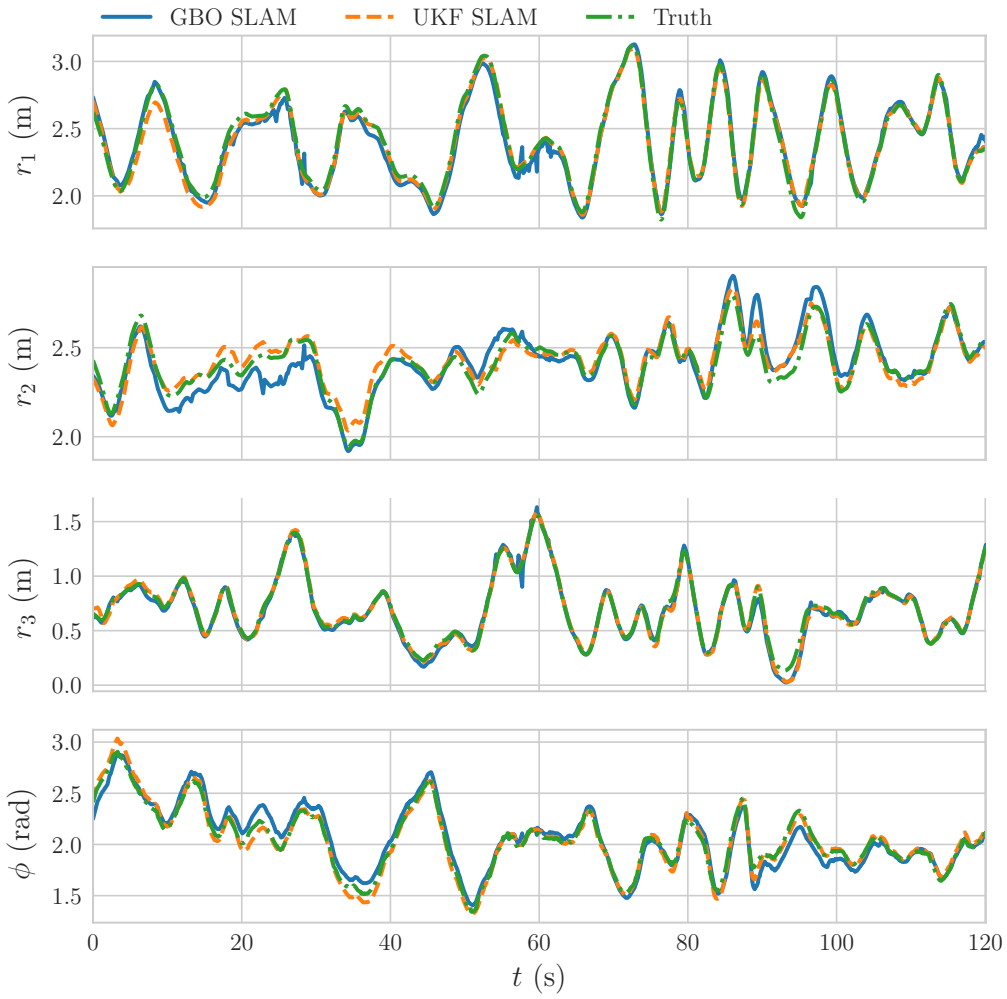


Figure 5.6: Time history of pose estimates and the true pose.

The performance of the proposed observer is shown in Fig. 5.6, Fig. 5.7, and Fig. 5.8, where the pose estimates, bias estimation error, and map estimation error is shown. Plotted are the first, second, and third components of $\mathbf{r}_a^{ba} = [\, r_1 \; r_2 \; r_3 \,]^\mathsf{T}$ as well as the attitude angle extracted from $\mathbf{C}_{ab}$ as $\phi = \cos^{-1}(\frac{1}{2}(\text{tr}(\mathbf{C}_{ab}) - 1))$. To compare

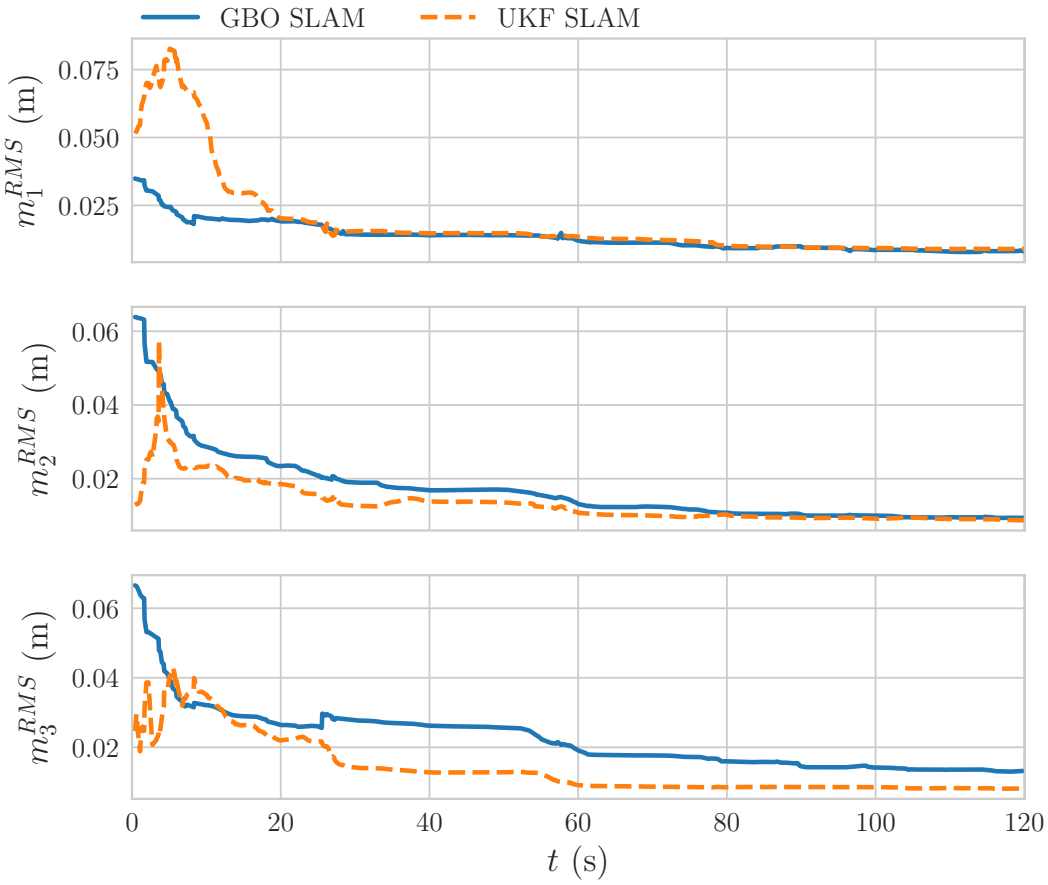


Figure 5.7: Map estimate RMS errors.

the pose estimate generated by the observer and the ground truth pose, the pose estimate is transformed by a best fit pose transformation, denoted $\mathbf{T}_{\text{fit}}$, computed using the final map estimate at time $t = 120$ (s) [85]. The RMS error associated with the first, second, and third components of all initialized landmark positions are also shown. Specifically, $m_i^{RMS} = \sqrt{\sum_{j\in\mathcal{I}}(m_{ai}^j - \bar{m}_{ai}^j)^2/N}$, $\forall i \in \{1,2,3\}$, where $\mathbf{m}_a^j = [m_{a1}^j \; m_{a2}^j \; m_{a3}^j \; 1]^\mathsf{T}$, $\bar{\mathbf{m}}_a^j = \mathbf{T}_{\text{fit}}^{-1}\hat{\mathbf{m}}_a^j$, $\mathcal{I}$ is the set of all initialized landmarks, and N is the total number of landmarks in $\mathcal{I}$.

For comparison purposes, an unscented Kalman Filter SLAM (UKF SLAM) algorithm is also implemented [86]. Due to the constant gain nature of the proposed observer, there exists a trade-off between convergence rate and steady-state error. Increasing the gains results in faster convergence but amplifies noise at steady-state. Therefore, the observer gains must be tuned appropriately. This is opposed to UKF SLAM where time-varying observer gains are generated from a covariance matrix

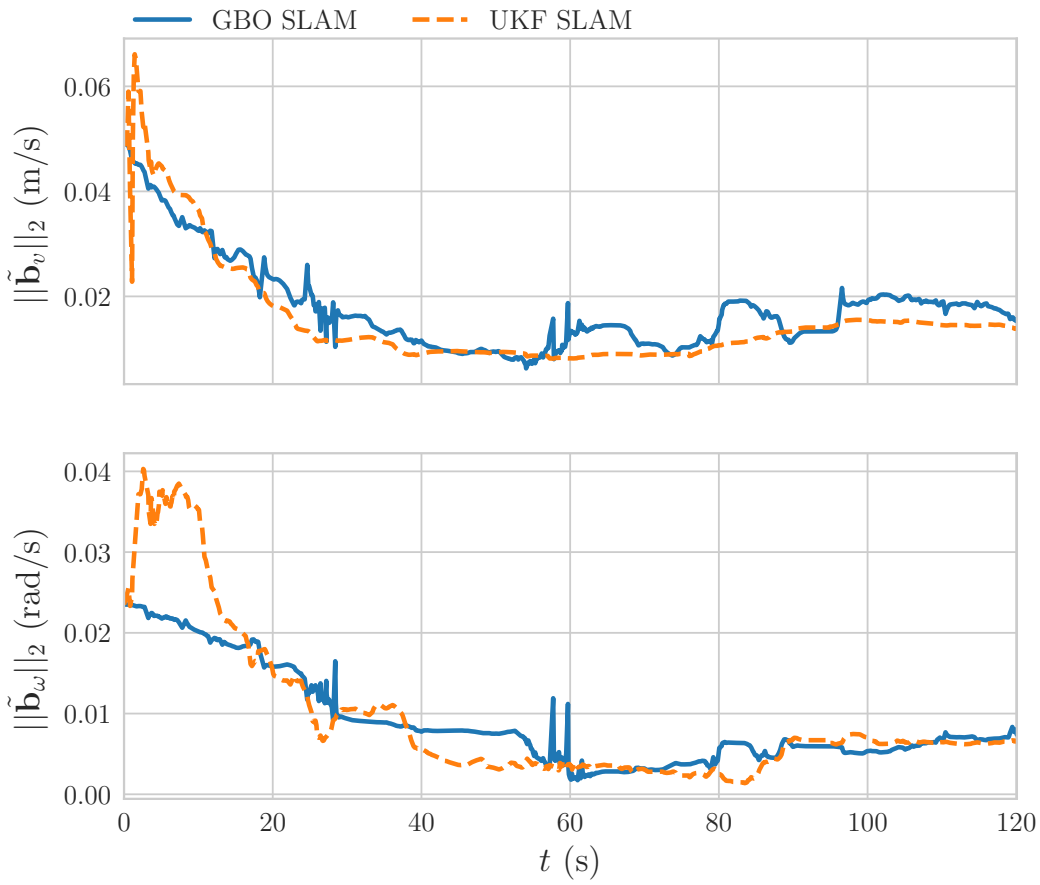


Figure 5.8: Time history of errors in the bias estimates.

that is propagated through time along with the state estimates. The UKF SLAM covariances are initially computed using truth and sensor data, then further tuned to achieve adequate performance. In this experiment, the GBO SLAM gains have been tuned to achieve similar steady-state performance as UKF SLAM, which can be seen in Fig. 5.6, Fig. 5.7, and Fig. 5.8. The results of this experiment are not meant to show that GBO SLAM is a superior algorithm to UKF SLAM, but rather that similar results can be achieved using the proposed algorithm when compared to traditional methods.

5.9 Closing Remarks

A gradient-based observer for SLAM has been presented. The observer evolves directly on the underlying Lie group. Stability results in Sec. 5.6 indicate that the pose

and map estimates converge to their true counterparts transformed by a constant pose. The case of estimating and negating bias corrupting measurements is also considered. Numerical and Experimental results demonstrate effective pose and map estimation. When compared with UKF SLAM the steady state pose and map errors of both UKF SLAM and GBO SLAM are comparable. Future work will involve including cross coupling between the pose and map innovations as well as time-varying observer gains.

Part III

Higher-Order Nonlinear Complementary Filtering

Chapter 6

The Higher-Order Nonlinear Complementary Filter

6.1 Introduction

Consider again the observer design problem for a general Lie group considered in Chapter 3. In that chapter, the gradient-based observer was discussed. As it will be shown in this chapter, the gradient-based observer is closely related to the classical complementary filter, a simple linear sensor fusion method. In fact, the gradient-based observer shares the same block diagram structure as a classical complementary filter, a fact pointed out for the $SO(3)$ case in [10]. The caveat, however, is that the gradient-based observer is only analogous to the specific case of a classical complementary filter with first order sensitivity and complementary sensitivity transfer functions. The first order nature nature of the gradient-based observer places restrictions on the effective bandwidth and roll-off rate of the observer.

In this chapter, a nonlinear Lie group observer is discussed that allows for higher-order filtering. The approach taken is similar to previous Lie group observers presented in the literature, including [14] and [15] as the innovation is related to the gradient of a cost function. However, while the gradient of the cost function appears directly as the innovation term in [14], in this chapter the innovation is based on the output of a linear time-invariant (LTI) system whose input is the gradient of a cost function resolved in a basis of the Lie algebra. Analogous to the classical complementary filter for any general sensitivity and complementary sensitivity transfer functions, the proposed method can be understood as a nonlinear complementary filter on a Lie group. The introduction of the LTI system in the proposed observer allows for more

general and complex higher-order filtering. Consequently, the proposed method completes the extension of linear complementary filters to nonlinear observer design on Lie groups. This is highly practical, as it allows for the targeting of specific frequency bandwidths in the velocity and partial state measurements. For example, applying a fast Fourier transform (FFT) to measurement data, the frequency content of the measurement noise can be identified and then mitigated by carefully constructing the LTI system that effectively acts as a prefilter on the exteroceptive measurements. Further, it is shown that by restricting the LTI system to the set of strictly positive real systems with feedthrough the strong stability properties of the nonlinear observer can be maintained. A disturbance observer that allows for the rejection of constant and harmonic disturbances in the group velocity measurements is also considered.

Passing nonlinear inputs through dynamic systems are considered in the context of rigid-body attitude estimation in [87] and [88], and for rigid-body attitude control in [89]. In [89] the nonlinear input is passed through a first-order dynamic system where the dynamic system is used to retain continuity in the case of discontinuities in the nonlinear input. In [87] and [88], a first-order dynamic system is employed and is derived via the Lagrange-D'Alembert principle and applied in continuous and discrete time. The filters in [87], [88], and [89] are first-order and consequently do not generalize to higher-order systems. Further, they were developed for the specific case of the Lie group $SO(3)$, while general Lie groups are considered in this chapter.

The main contribution of this chapter is the design of the observer, which allows for greater design freedom compared to similar observers and can, with appropriate selection of the LTI system, result in enhanced rejection of measurement noise. It is shown that, provided the linear system is composed of a strictly positive real part and feedthrough, the proposed observer is locally asymptotically stable about a desired equilibrium point.

6.2 Preliminaries

As in Chapter 3, let G denote a finite-dimensional Lie group and $\mathfrak{g}$ its associated Lie algebra. Let the set $B = \{b_1, \ldots, b_n\}$, where $b_1, \ldots, b_n \in \mathfrak{g}$, be a basis of $\mathfrak{g}$. Then, for any $a \in \mathfrak{g}$, a may be written as $a = S(\mathbf{a})$ where $\mathbf{a} = [\ a_1, \ldots, a_n\ ]^\mathsf{T} \in \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathfrak{g}$ is such that $S(\mathbf{a}) = \sum_{i=1}^{n} a_i b_i$. A basis of $T_X G$ may be found by right translation of B by X such that $B_X = \{b_1 X, \ldots, b_n X\}$ is a basis of $T_X G$. To simplify the results included in this chapter the basis B is assumed to be orthonormal, however, similar results follow for any arbitrary basis.

An inner product on $\mathfrak{g}$ is denoted $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ with associated norm $|| \cdot ||_{\mathfrak{g}} = \sqrt{\langle \cdot, \cdot \rangle}$. In this chapter, when the Riemannian metric is assumed to be right-invariant it will also be assumed that the Riemannian metric is induced by the inner product on $\mathfrak{g}$ by the relationship

$$\begin{aligned} \langle V(X), U(X) \rangle_X &= \langle V(X)X^{-1}, U(X)X^{-1} \rangle \\ &= \langle v, u \rangle \\ &= \mathbf{v}^\mathsf{T}\mathbf{u} \end{aligned} \tag{6.1}$$

for all $X \in G$, and vector fields $V(X) = vX$ and $U(X) = uX$, where $v = S(\mathbf{v}), u = S(\mathbf{u}) \in \mathfrak{g}$ [17]. When the Riemannian metric is assumed to be bi-invariant, it is also assumed that the inner product on $\mathfrak{g}$ gives rise to a bi-invariant Riemannian metric by the same relationship.

6.3 Observer Design

Consider again the problem of observer design on a Lie group G discussed in Chapter 3, where it was desired to build an observer for the system (3.2) with measurements given in (3.3) and (3.8). For convenience, (3.2), (3.3), and (3.8) are repeated here as

$$\dot{X}(t) = X(t)v(t), \tag{6.2}$$

and

$$v_y(t) = v(t) + d(t), \tag{6.3a}$$

$$y = h(X(t), N(t), \bar{y}). \tag{6.3b}$$

6.3.1 Review of the Linear Complementary Filter

To motivate the design of the observer presented in this chapter, first consider the analogous system to (6.2) on $\mathbb{R}$ given by

$$\dot{x}(t) = v(t),$$

where $x \in \mathbb{R}$ and $v \in \mathbb{R}$ is some time-dependent exogenous signal. Suppose it is desired to build a filter to estimate the state x from measurements of x and v given

by

$$\begin{aligned} y(t) &= x(t) + n(t), \\ v_y(t) &= v(t) + d(t), \end{aligned}$$

where $n, d \in \mathbb{R}$ are, respectively, the noise associated with y and v_y. The complementary filter is a simple method to fuse the measurements of $x(\cdot)$ and $v(\cdot)$, and is particularly effective when y and v_y have complementary noise characteristics [90, 91]. Expressed in the frequency domain, the classical complementary filter is given by

$$\begin{aligned} \hat{x}(s) &= S(s)\frac{v_y(s)}{s} + T(s)y(s) \\ &= x(s) + S(s)\frac{w(s)}{s} + T(s)n(s) \end{aligned} \tag{6.4}$$

where $S(s)$ and $T(s)$ are the sensitivity and complementary sensitivity functions of the closed-loop system such that $S(s)+T(s) = 1$ [10, 90, 91]. Letting $S(s) = s/(s+H(s))$ and $T(s) = H(s)/(s + H(s))$ for some transfer function $H(s)$, (6.4) can be expressed as

$$\hat{x}(s) = x(s) + \frac{s}{s + H(s)}\frac{d(s)}{s} + \frac{H(s)}{s + H(s)}n(s). \tag{6.5}$$

In state-space form (6.5) can be written as

$$\dot{\hat{x}} = v_y - u \tag{6.6a}$$

$$\dot{\mathbf{x}}_f = \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f(\hat{x} - y) \tag{6.6b}$$

$$u = \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f(\hat{x} - y), \tag{6.6c}$$

where $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f, \mathbf{D}_f)$ forms a minimal state-space realization of $H(s)$, $\mathbf{x}_f$ is the state associated with $H(s)$, and $u(s) = H(s)(\hat{x}(s) - y(s))$. Defining the error $e = \hat{x} - x$, the error dynamics associated with (6.6) are

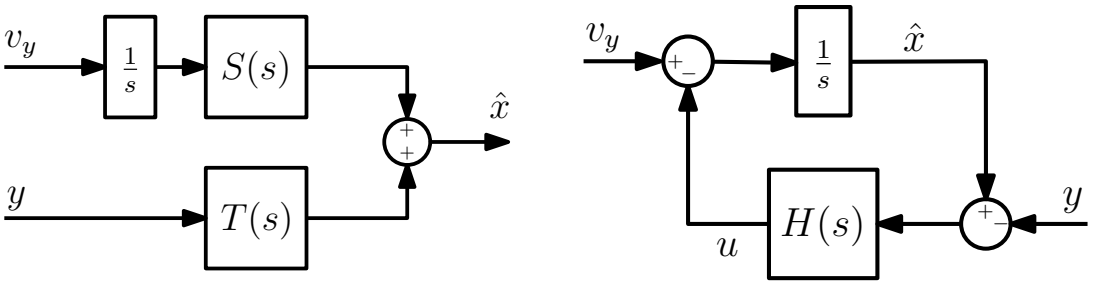
$$\dot{e} = d - u \tag{6.7a}$$

$$\dot{\mathbf{x}}_f = \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f(e - n) \tag{6.7b}$$

$$u = \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f(e - n). \tag{6.7c}$$

The linear complementary filter is summarized in Fig. 6.1, where the block diagrams associated with (6.4) and (6.6) are shown.

Classical control methods can be used to design $H(s)$ such that $T(s)$ and $S(s)$



(a) The state estimate is a fusion of measurements v_y and y.

(b) Feedback structure of the complementary filter.

Figure 6.1: Equivalent block diagrams of the linear complementary filter. The complementary filter can be alternatively thought of as an open-loop sensor fusion algorithm or as a feedback control system.

have desirable properties. When $n(s)$ is comprised of high frequency noise and $d(s)/s$ is comprised of low frequency noise, $H(s)$ is designed such that $T(s)$ and $S(s)$ are low-pass and high-pass filters, respectively. A simple method to accomplish this is to let $H(s) = k$, where $k \in (0, \infty)$. Then, $T(s)$ and $S(s)$ respectively become first order low and high-pass filters with cutoff frequencies of k (rad/s).

6.3.2 Nonlinear Complementary Filtering

Motivated by (6.6), the Lie group observers proposed in this chapter take the form of either

$$\dot{\hat{X}} = \hat{X} v_y - u_R \hat{X}, \tag{6.8a}$$

$$\mathbf{u}_R = \boldsymbol{\mathcal{H}} \mathbf{e}_R, \tag{6.8b}$$

or

$$\dot{\hat{X}} = \hat{X} v_y - \hat{X} u_L, \tag{6.9a}$$

$$\mathbf{u}_L = \boldsymbol{\mathcal{H}} \mathbf{e}_L, \tag{6.9b}$$

where

$$e_R = T_{\hat{X}} R_{\hat{X}^{-1}}(\nabla_{\hat{X}} f_y(\hat{X}, y)), \quad e_L = T_{\hat{X}} L_{\hat{X}^{-1}}(\nabla_{\hat{X}} f_y(\hat{X}, y)), \tag{6.10}$$

$\boldsymbol{\mathcal{H}}$ is a linear operator, and $\mathbf{e}_R, \mathbf{e}_L, \mathbf{u}_R, \mathbf{u}_L \in \mathbb{R}^n$ are respectively the representations of e_R, e_L, u_R, and u_L in the basis B. The observer in (6.8) will be referred to as the right observer while the observer in (6.12) will be referred to as the left observer. In this dissertation, the linear operator $\boldsymbol{\mathcal{H}}$ is assumed to be linear time-invariant (LTI). The

LTI system can be represented as the transfer matrix $\mathbf{H}(s) = \mathbf{C}_f(s\mathbf{1} - \mathbf{A}_f)^{-1}\mathbf{B}_f + \mathbf{D}_f$ with associated state $\mathbf{x}_f \in \mathbb{R}^{n_f}$. In this case, the proposed observers take the form

$$\begin{aligned}
\dot{\hat{X}} &= \hat{X}v_y - S(\mathbf{u}_R)\hat{X}, && (6.11a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && (6.11b)\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && (6.11c)\\
\mathbf{e}_R &= S^{-1}(T_{\hat{X}}R_{\hat{X}^{-1}}(\nabla_{\hat{X}}f_y(\hat{X}, y))), && (6.11d)
\end{aligned}$$

or

$$\begin{aligned}
\dot{\hat{X}} &= \hat{X}v_y - \hat{X}S(\mathbf{u}_L), && (6.12a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && (6.12b)\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && (6.12c)\\
\mathbf{e}_L &= S^{-1}(T_{\hat{X}}L_{\hat{X}^{-1}}(\nabla_{\hat{X}}f_y(\hat{X}, y))). && (6.12d)
\end{aligned}$$

The proposed observers are composed of two coupled ordinary differential equations. The first, (6.11a) and (6.12a), evolves directly on the underlying Lie group G, while the second, (6.11b) and (6.12b), is a linear system evolving on $\mathbb{R}^{n_f}$. Taken on its own, (6.11a) shares the same structure as previous Lie group observers proposed in the literature, including [14, 16], in that it is composed of two terms, the first of which copies the nonlinear system dynamics of (3.2) and the second is an innovation term that serves to drive the state estimate towards the true state. In fact, taking $\mathbf{H}(s) = k\mathbf{1}$, $k \in (0, \infty)$, and neglecting partial state measurement noise, the proposed observer reduces to

$$\dot{\hat{X}} = \hat{X}v_y - k\nabla_{\hat{X}}f_{\mathcal{I}}(\hat{X}, X), \tag{6.13}$$

the left gradient observer proposed in [14] and discussed in Chapter 3.

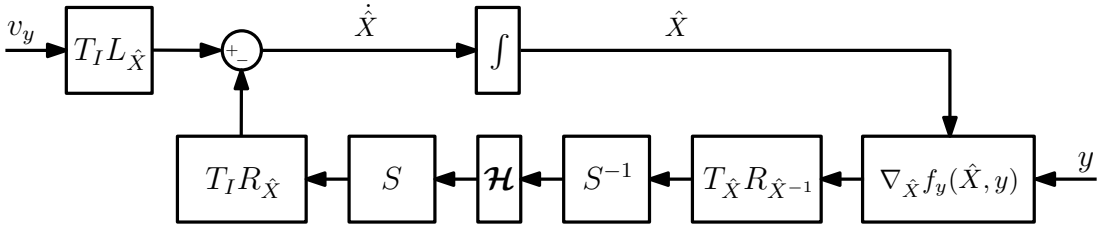


Figure 6.2: Block diagram of the proposed right observer with group velocity measurement v_y and partial state measurements y.

Noting the similarities in structure between the classical complementary filter (6.6)

to the proposed observers (6.11)-(6.12), the proposed methods can be understood as nonlinear complementary filters on the Lie group G. The similarities are even more apparent by comparing block diagrams of the linear complementary filter in Fig. 6.1b with the block diagram of the proposed right observer shown in Fig. 6.2. The similarities between Lie group observers of the form (6.13) and linear complementary filters with a constant transfer function $H(s) = k$ was first noted for the case of the Lie group $SO(3)$ in [10]. The proposed observer, however, is analogous to a classical complementary filter on $\mathbb{R}$ for any general transfer function $H(s)$, rather than strictly for constant $H(s)$. Therefore, the introduction of the linear system $\mathbf{H}(s)$ in (6.11)-(6.12) completes the extension of linear complementary filters, with any general transfer function $H(s)$, to nonlinear complementary filters on Lie groups. In practice, $\mathbf{H}(s)$ allows for greater freedom in the design of sensitivity and complementary sensitivity transfer functions when (6.11) is linearized. A constant transfer function only allows for simple first-order low- and high-pass filtering, while higher-order filtering can be accomplished with the appropriate selection of $\mathbf{H}(s)$. This enhanced design freedom can be exploited to better reject measurement noise and improve performance of the nonlinear observer. It is shown in Sec. 6.4 that the strong stability properties typical of nonlinear Lie group observers can be maintained even with the introduction of $\mathbf{H}(s)$.

6.3.3 Right Observer Error Dynamics

To analyze the stability of the right observer it will be helpful to determine the dynamics of $(\tilde{X}_R, \mathbf{x}_f)$. As is the case in [14], left-invariant system dynamics along with the right-invariance of the chosen Riemannian metric and cost function yield autonomous error dynamics for the right observer. The autonomy of the error dynamics in the absence of noise is established in the following proposition.

Proposition 6.1. *Consider trajectories of* $(\hat{X}, \mathbf{x}_f)$ *under the right observer in* (6.11). *Let* $f_N : G \times G \to \mathbb{R}$ *be a right-invariant cost function and take a right-invariant Riemannian metric on* G. *Assume that group velocity measurements and partial state measurements are available as in* (6.3). *Then, dynamics associated with* $(\tilde{X}_R, \mathbf{x}_f)$ *are*

given by

$$\begin{aligned}
\dot{\tilde{X}}_R &= \tilde{X}_R \mathrm{Ad}_X(d) - S(\mathbf{u}_R)\tilde{X}_R, && (6.14a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && (6.14b)\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && (6.14c)\\
\mathbf{e}_R &= S^{-1}(T_{\tilde{X}_R} R_{\tilde{X}_R^{-1}} \nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I)). && (6.14d)
\end{aligned}$$

Further, assuming that there is no noise associated with the group velocity and partial state measurements such that $d = 0$ and $N = \mathcal{I}$, the dynamics associated with $(\tilde{X}_R, \mathbf{x}_f)$ are autonomous and are given by

$$\begin{aligned}
\dot{\tilde{X}}_R &= -S(\mathbf{u}_R)\tilde{X}_R, && (6.15a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && (6.15b)\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && (6.15c)\\
\mathbf{e}_R &= S^{-1}(T_{\tilde{X}_R} R_{\tilde{X}_R^{-1}} \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I)). && (6.15d)
\end{aligned}$$

Proof The proof follows in a similar manner to the proof of Theorem 17 in [14]. By Lemma 3.1, the expression for $\dot{\tilde{X}}_R$ satisfies

$$\begin{aligned}
\dot{\tilde{X}}_R &= T_{\hat{X}} R_{X^{-1}}(\hat{X}d - u_R\hat{X})\\
&= T_{\hat{X}} R_{X^{-1}}(\hat{X}d) - T_{\hat{X}} R_{X^{-1}}(u_R\hat{X})\\
&= T_{\hat{X}} R_{X^{-1}}(\hat{X}d) - u_R\tilde{X}.
\end{aligned}$$

Recall from (6.11d), $\mathbf{e}_R = S^{-1}(T_{\hat{X}} R_{\hat{X}^{-1}}(\nabla_{\hat{X}} f_y(\hat{X}, y)))$, or alternatively $\nabla_{\hat{X}} f_N(\hat{X}, X) = S(\mathbf{e}_R)\hat{X}$. By Lemma 3.6, it follows that $\nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I) = T_{\hat{X}} R_{X^{-1}}(\nabla_{\hat{X}} f_N(\hat{X}, X))$, and therefore

$$\begin{aligned}
\nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I) &= T_{\hat{X}} R_{X^{-1}}(S(\mathbf{e}_R)\hat{X})\\
&= S(\mathbf{e}_R)\tilde{X}_R,
\end{aligned}$$

which implies that $\mathbf{e}_R = S^{-1}(T_{\tilde{X}_R} R_{\tilde{X}_R^{-1}} \nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I))$. This proves that the error dynamics satisfy (6.14). Equation (6.15) follows simply by setting $d = 0$ and $N = \mathcal{I}$. The error dynamics are autonomous in this case as $\mathbf{e}_R$ depends only on $\tilde{X}_R$. □

6.3.4 Left Observer Error Dynamics

To prove the stability properties of the left observer the error dynamics associated with $(\tilde{X}_R, \mathbf{x}_f)$ will be used rather than $(\tilde{X}_L, \mathbf{x}_f)$. This is done to simplify the stability proofs. However, in later chapters the error dynamics associated with $(\tilde{X}_L, \mathbf{x}_f)$ will be used for design purposes. Therefore, in the following proposition, the error dynamics under (6.12) for both $(\tilde{X}_R, \mathbf{x}_f)$ and $(\tilde{X}_L, \mathbf{x}_f)$ are determined.

Proposition 6.2. *Consider trajectories of $(\hat{X}, \mathbf{x}_f)$ under the left observer in* (6.12)*. Let $f_N : G \times G \to \mathbb{R}$ be a bi-invariant cost function and take a bi-invariant Riemannian metric on G. Assume that group velocity measurements and partial state measurements are available as in* (6.3)*. Then, the error dynamics associated with $(\tilde{X}_L, \mathbf{x}_f)$ are given by*

$$\begin{aligned}
\dot{\tilde{X}}_L &= -v\tilde{X}_L + \tilde{X}_L v + \tilde{X}_L(d - u_L), && \text{(6.16a)}\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && \text{(6.16b)}\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && \text{(6.16c)}\\
\mathbf{e}_L &= S^{-1}(T_{\tilde{X}_L}L_{\tilde{X}_L^{-1}}(\nabla_{\tilde{X}_L} f_N(\tilde{X}_L, I))), && \text{(6.16d)}
\end{aligned}$$

and the error dynamics associated with $(\tilde{X}_R, \mathbf{x}_f)$ are given by

$$\begin{aligned}
\dot{\tilde{X}}_R &= \tilde{X}_R\mathrm{Ad}_X(d) - \mathrm{Ad}_{\hat{X}}(u_L), && \text{(6.17a)}\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && \text{(6.17b)}\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && \text{(6.17c)}\\
\mathbf{e}_L &= S^{-1}(\mathrm{Ad}_{\hat{X}^{-1}}(T_{\tilde{X}_R}R_{\tilde{X}_R^{-1}}(\nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I)))), && \text{(6.17d)}
\end{aligned}$$

Further, assuming that there is no noise associated with the group velocity and partial state measurements such that $d = 0$ and $N = \mathcal{I}$. Then, the dynamics associated with $(\tilde{X}_L, \mathbf{x}_f)$ are nonautonomous and are given by

$$\begin{aligned}
\dot{\tilde{X}}_L &= -v\tilde{X}_L + \tilde{X}_L v - \tilde{X}_L u_L, && \text{(6.18a)}\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && \text{(6.18b)}\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && \text{(6.18c)}\\
\mathbf{e}_L &= S^{-1}(T_{\tilde{X}_L}L_{\tilde{X}_L^{-1}}(\nabla_{\tilde{X}_L} f_{\mathcal{I}}(\tilde{X}_L, I))), && \text{(6.18d)}
\end{aligned}$$

and the error dynamics associated with $(\tilde{X}_R, \mathbf{x}_f)$ *are given by*

$$\begin{aligned}
\dot{\tilde{X}}_R &= -\mathrm{Ad}_{\hat{X}}(u_L)\tilde{X}_R, && (6.19a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && (6.19b)\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && (6.19c)\\
\mathbf{e}_L &= \mathbf{e}_L = S^{-1}(\mathrm{Ad}_{\hat{X}^{-1}}(T_{\tilde{X}_R}R_{\tilde{X}_R^{-1}}(\nabla_{\tilde{X}_R}f_{\mathcal{I}}(\tilde{X}_R, I)))), && (6.19d)
\end{aligned}$$

Proof. Again, the proof follows in a similar manner to the proof of Theorem 17 in [14]. To show (6.16), the time derivative of $\tilde{X}_L = X^{-1}\hat{X}$ will be taken using Lemma 3.1. By Lemma 3.1,

$$\begin{aligned}
\dot{\tilde{X}}_L &= -T_{\hat{X}^{-1}}R_{\hat{X}}T_I L_{X^{-1}}T_X R_{X^{-1}}\dot{X} + T_{\hat{X}}L_{X^{-1}}\dot{\hat{X}}\\
&= -X^{-1}XvX^{-1}\hat{X} + X^{-1}(\hat{X}(v+d) - \hat{X}u_L)\\
&= -v\tilde{X}_L + \tilde{X}_L v + \tilde{X}_L(d - u_L),
\end{aligned}$$

which shows (6.16). Recall from (6.12d), $\mathbf{e}_L = S^{-1}(T_{\hat{X}}L_{\hat{X}^{-1}}(\nabla_{\hat{X}}f_y(\hat{X}, y)))$, or alternatively $\nabla_{\hat{X}}f_N(\hat{X}, X) = \hat{X}S(\mathbf{e}_L)$. As the cost function and Riemannian metric are bi-invariant it follows that they are left-invariant and by Lemma 3.6, $\nabla_{\tilde{X}_L}f_N(\tilde{X}_L, I) = T_{\hat{X}}L_{X^{-1}}(\nabla_{\hat{X}}f_N(\hat{X}, X))$, and therefore

$$\begin{aligned}
\nabla_{\tilde{X}_L}f_N(\tilde{X}_R, I) &= T_{\hat{X}}L_{X^{-1}}(\hat{X}S(\mathbf{e}_L))\\
&= \tilde{X}_L S(\mathbf{e}_L),
\end{aligned}$$

which implies that $\mathbf{e}_L = S^{-1}(T_{\tilde{X}_L}L_{\tilde{X}_L^{-1}}\nabla_{\tilde{X}_L}f_N(\tilde{X}_L, I))$. This proves that the error dynamics satisfy (6.16).

To show (6.16), Lemma 3.1 is applied again to yield $\dot{\tilde{X}}_R = \tilde{X}_R\mathrm{Ad}_X(d) - \mathrm{Ad}_{\hat{X}}(u_L)$. As the cost function and Riemannian metric are bi-invariant it follows that they are also right-invariant and by Lemma 3.6 $\nabla_{\tilde{X}_R}f_N(\tilde{X}_R, I) = T_{\hat{X}}R_{X^{-1}}(\nabla_{\hat{X}}f_N(\hat{X}, X))$. This implies that

$$\begin{aligned}
\nabla_{\tilde{X}_R}f_N(\tilde{X}_R, I) &= T_{\hat{X}}R_{X^{-1}}(\hat{X}e_L)\\
&= \hat{X}e_L X^{-1}\\
&= \mathrm{Ad}_{\hat{X}}(e_L)\tilde{X}_R,
\end{aligned}$$

which implies that $\mathbf{e}_L = S^{-1}(\mathrm{Ad}_{\hat{X}^{-1}}(T_{\tilde{X}_R}R_{\tilde{X}_R^{-1}}(\nabla_{\tilde{X}_R}f_N(\tilde{X}_R, I))))$. Equations (6.18) and (6.18) follow by setting $d = 0$ and $N = \mathcal{I}$. □

6.4 Stability Results

In the stability results that follow restrictions will be made on the cost function, f, as well as the linear system $\mathbf{H}(s)$. In particular, the cost function will be restricted to the set of right- or bi-invariant error functions, as defined below, and the linear system is restricted to the set of strictly positive real systems with feedthrough.

Definition 6.3 (Error function [17])**.** Consider a smooth symmetric function $\psi : G \times G \to \mathbb{R}$ and define $\psi_X : G \to \mathbb{R}$, where $\psi_X(Y) = \psi(X, Y)$ for all $X, Y \in G$. The function $\psi : G \times G \to \mathbb{R}$ is an *error function* about $X \in G$ if ψ_X is smooth, proper, bounded from below, and ψ_X satisfies

(i) $\psi_X(X) = 0$,

(ii) $\nabla\psi_X(X) = 0$,

(iii) $H_{\psi_X(X)}$ is positive definite.

The properties of an error function are well established in [17], where the error function is labeled a "tracking error function". A method for constructing error functions based on single variable cost functions on the output spaces is proposed in [15]. Another method for finding right invariant cost functions is discussed in [14].

Definition 6.4 (Strictly Positive Real (SPR) Transfer Matrix [92])**.** A real, rational, strictly proper transfer matrix $\mathbf{H}_{\text{spr}}(s)$ of the complex variable s is SPR if

1. $\mathbf{H}_{\text{spr}}(s)$ is real for all real s and all elements of $\mathbf{H}_{\text{spr}}(s)$ are analytic in $Re\{s\} \geq 0$,
2. $\mathbf{H}_{\text{spr}}(j\omega) + \mathbf{H}_{\text{spr}}^{\mathsf{H}}(j\omega) > \mathbf{0}\ \forall\omega \in (-\infty, \infty)$,
3. $\lim_{\omega\to\infty} \omega^2\{\mathbf{H}_{\text{spr}}(j\omega) + \mathbf{H}_{\text{spr}}^{\mathsf{H}}(j\omega)\} > \mathbf{0}$.

Lemma 6.5 (Kalman-Yakubovich-Popov (KYP) Lemma [93])**.** *Consider the LTI system*

$$\begin{aligned} \dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{u}_f, \\ \mathbf{y}_f &= \mathbf{C}_f\mathbf{x}_f, \end{aligned}$$

where $\mathbf{x}_f \in \mathbb{R}^{n_f}$, $\mathbf{u}_f, \mathbf{y}_f \in \mathbb{R}^{m_f}$, *and* $\mathbf{A}_f$, $\mathbf{B}_f$, *and* $\mathbf{C}_f$ *are appropriately dimensioned real matrices that form a minimal state-space realization. Moreover, assume that* $\mathbf{A}_f$

is Hurwitz. Then, the system is strictly positive real (SPR) if and only if there exists symmetric positive definite matrices $\mathbf{P}_f, \mathbf{Q}_f \in \mathbb{R}^{n_f \times n_f}$ *such that*

$$\mathbf{P}_f \mathbf{A}_f + \mathbf{A}_f^\mathsf{T} \mathbf{P}_f = -\mathbf{Q}_f \tag{6.20a}$$

$$\mathbf{P}_f \mathbf{B}_f = \mathbf{C}_f^\mathsf{T}. \tag{6.20b}$$

Corollary 6.6. *Consider an LTI system with minimal state-space realization given by*

$$\dot{\mathbf{x}}_f = \mathbf{A}_f \mathbf{x}_f + \mathbf{B}_f \mathbf{u}_f, \tag{6.21a}$$

$$\mathbf{y}_f = \mathbf{C}_f \mathbf{x}_f + \mathbf{D}_f \mathbf{u}_f, \tag{6.21b}$$

and define $\mathcal{L}(\mathbf{x}_f) = \frac{1}{2}\mathbf{x}_f^\mathsf{T} \mathbf{P}_f \mathbf{x}_f$. *If matrices* $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f)$ *satisfy the KYP equations* (6.20) *then, regardless of the choice of* $\mathbf{D}_f$, *the time derivative of* $\mathcal{L}$ *is given by*

$$\dot{\mathcal{L}}(\mathbf{x}_f) = -\tfrac{1}{2}\mathbf{x}_f^\mathsf{T} \mathbf{Q}_f \mathbf{x}_f + \mathbf{u}_f^\mathsf{T} \mathbf{y}_f - \mathbf{u}_f^\mathsf{T} \mathbf{D}_f \mathbf{u}_f.$$

Proof The proof follows directly from equations (6.20) and (6.21). □

Given Definition 6.3 and Definition 6.4 it is now possible to present the main results of this section. As in [16] we require the existence of a faithful representation of the Lie group G as a matrix Lie group.

6.4.1 Stability of the Right Observer

The following corollary will be useful in the proof of the theorem that follows.

Corollary 6.7. *Consider the gradient vector field* $\nabla f_{\mathcal{I}}(\tilde{X}_R, I) = S(\mathbf{e}_R)\tilde{X}_R$, *where the gradient is taken with respect to a right-invariant Riemannian metric, and let* $\dot{\tilde{X}}_R = S(\mathbf{q})\tilde{X}_R$, *where* $\mathbf{q} \in \mathbb{R}^n$. *Then, the derivative with respect to time of* $\mathbf{e}_R$ *is given by*

$$\dot{\mathbf{e}}_R = \mathbf{H}(\tilde{X}_R)\mathbf{q} - \boldsymbol{\xi},$$

where $\mathbf{H}(\tilde{X}_R)$ *is the matrix representation in basis* $B_{\tilde{X}_R}$ *of the Riemannian Hessian operator* $H_{f_{\mathcal{I}}(\tilde{X}_R, I)}(\cdot)$ *at point* $\tilde{X}$, *and* $\boldsymbol{\xi} \in \mathbb{R}^n$. *Moreover, there exists a finite constant* $m < \infty$ *such that* $\boldsymbol{\xi}$ *satisfies* $||\boldsymbol{\xi}||_2 \leq m||\mathbf{q}||_2||\mathbf{e}_R||_2$.

Proof Let $\tilde{X}_R$ be a trajectory under the ordinary differential equation $\dot{\tilde{X}}_R = S(\mathbf{q})\tilde{X}_R$, and consider the gradient vector field $\nabla f_{\mathcal{I}}(\tilde{X}_R, I)$. Let Γ be a vector field along the curve $\tilde{X}_R$ such that $\Gamma(t) = \nabla f_{\mathcal{I}}(\tilde{X}_R(t), I)$. The covariant derivative of Γ

along $\tilde{X}_R$ is given by [17, p. 139],

$$\frac{\mathrm{D}\Gamma}{\mathrm{d}t} = S(\dot{\mathbf{e}}_R)\tilde{X}_R + Q(S(\mathbf{q}), S(\mathbf{e}_R))\tilde{X}_R, \tag{6.22}$$

where $Q : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is the unique bilinear mapping associated with the Levi-Civita connection such that for any two right invariant vector fields $V(X) = vX$ and $U(X) = uX$, $\nabla_V U = Q(v, u)X$. By definition the covariant derivative satisfies

$$\frac{\mathrm{D}\Gamma}{\mathrm{d}t} = \nabla_{\dot{\tilde{X}}_R} \nabla f_{\mathcal{I}}(\tilde{X}_R, I),$$

and from the definition of the Hessian operator in (2.8) it follows that

$$\frac{\mathrm{D}\Gamma}{\mathrm{d}t} = H_{f_{\mathcal{I}}(\tilde{X}_R, I)}(\dot{\tilde{X}}_R). \tag{6.23}$$

Combining (6.22) and (6.23) yields

$$S(\dot{\mathbf{e}}_R)\tilde{X}_R = H_{f_{\mathcal{I}}(\tilde{X}_R, I)}(\dot{\tilde{X}}_R) - Q(S(\mathbf{q}), S(\mathbf{e}_R))\tilde{X}_R, \tag{6.24}$$

Resolving (6.24) in basis $B_{\tilde{X}_R}$ yields

$$\dot{\mathbf{e}}_R = \mathbf{H}(\tilde{X}_R)\mathbf{q} - \boldsymbol{\xi},$$

where $S(\boldsymbol{\xi}) = Q(S(\mathbf{q}), S(\mathbf{e}_R))$. Recall that the Riemannian manifold under consideration, namely $(G, \langle \cdot, \cdot \rangle_X)$, is a smooth Riemannian manifold. This implies that the Levi-Civita connection is continuous [17, p.115 Theorem 3.104]. This further implies that the bilinear map $Q(\cdot, \cdot)$ is continuous and consequently $Q(\cdot, \cdot)$ is bounded, which is to say that there exists $m < \infty$ such that $||Q(a, b)||_{\mathfrak{g}} \leq m||a||_{\mathfrak{g}}||b||_{\mathfrak{g}}$ for all $a, b \in \mathfrak{g}$. Thus, $||S(\boldsymbol{\xi})||_{\mathfrak{g}} \leq m||S(\mathbf{q})||_{\mathfrak{g}}||S(\mathbf{e}_R)||_{\mathfrak{g}}$, which implies that $||\boldsymbol{\xi}||_2 \leq m||\mathbf{q}||_2||\mathbf{e}_R||_2$. □

Theorem 6.8. *Consider trajectories of* $(\tilde{X}_R, \mathbf{x}_f)$ *under* (6.15). *Let* $f_{\mathcal{I}}$ *be a right-invariant error function about* X *and take a right-invariant Riemannian metric. Let* $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f)$ *be SPR where* $\mathbf{B}_f$ *has full rank, and let* $\mathbf{D}_f \geq 0$. *Assume the following:*

(A1) there is no noise associated with the group velocity measurement or the partial state measurements such that $d = 0$ *and* $N = \mathcal{I}$;

(A2) there exists a faithful representation of the Lie group G *as a matrix Lie group.*

Define $L = \sup\{c \in \mathbb{R} \mid \tilde{X}_R \in \Omega_c \setminus \{I\} \implies \mathbf{e}_R \neq \mathbf{0}\}$, *where* $\Omega_c = \{\tilde{X}_R \in G \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq c\}$. *Then the following statements hold:*

(i) the equilibrium point $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ *is locally asymptotically stable;*

(ii) trajectories of $(\mathbf{e}_R, \mathbf{x}_f)$ *exponentially approach* $(\mathbf{0}, \mathbf{0})$ *and* $\tilde{X}_R$ *asymptotically approaches* I *for all initial conditions satisfying* $V_1(\tilde{X}_R(0), \mathbf{x}_f(0)) < L$*, where* V_1 *is defined in* (6.29).

Proof. First, item (i) will be proved. Consider the Lyapunov function candidate

$$V_1(\tilde{X}_R, \mathbf{x}_f) = f_{\mathcal{I}}(\tilde{X}_R, I) + \mathcal{L}(\mathbf{x}_f). \tag{6.25}$$

The derivative with respect to time of V_1 is

$$\begin{aligned} \dot{V}_1(\tilde{X}_R, \mathbf{x}_f) &= \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) \\ &= -\langle S(\mathbf{e}_R)\tilde{X}_R, S(\mathbf{u})\tilde{X}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) \\ &= -\langle S(\mathbf{e}_R), S(\mathbf{u}_R) \rangle + \dot{\mathcal{L}}(\mathbf{x}_f) \\ &= -\mathbf{e}_R^{\mathsf{T}} \mathbf{u}_R + \dot{\mathcal{L}}(\mathbf{x}_f). \end{aligned}$$

By application of Corollary 6.6, $\dot{V}_1$ is given by

$$\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) = -\mathbf{e}_R^{\mathsf{T}} \mathbf{u}_R - \tfrac{1}{2} \mathbf{x}_f^{\mathsf{T}} \mathbf{Q}_f \mathbf{x}_f + \mathbf{e}_R^{\mathsf{T}} \mathbf{u}_R - \mathbf{e}_R^{\mathsf{T}} \mathbf{D}_f \mathbf{e}_R.$$

Consequently, $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) = -\frac{1}{2} \mathbf{x}_f^{\mathsf{T}} \mathbf{Q}_f \mathbf{x}_f - \mathbf{e}_R^{\mathsf{T}} \mathbf{D}_f \mathbf{e}_R$. By assumption $\mathbf{D}_f \geq 0$, which implies $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \leq -\frac{1}{2} \mathbf{x}_f^{\mathsf{T}} \mathbf{Q}_f \mathbf{x}_f$ and thus $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \leq 0$ and $V_1(\tilde{X}_R(t), \mathbf{x}_f(t)) \leq V_1(\tilde{X}_R(0), \mathbf{x}_f(0))$ for all $t \geq 0$. By Remark 11.11 of [17], $f_{\mathcal{I}}$ is locally positive definite and $\tilde{X} = I$ is an isolated critical point of $f_{\mathcal{I}}$. By assumption, there exists a faithful representation of G as a matrix Lie group. This implies that there exists $m > 0$ and a mapping $\mathbf{\Phi} : G \to GL(m)$ such that $\mathbf{\Phi}(G)$ is a matrix Lie group [16]. Following the proof of Theorem 5.1 in [16], this implies that there exists a set $B_r = \{\tilde{X}_R \in G \mid d(\tilde{X}_R) \leq r\}$ about $\tilde{X}_R = I$, where $d(\tilde{X}_R) = ||\mathbf{1} - \mathbf{\Phi}(\tilde{X}_R)||_{\mathsf{F}}$, such that for all $\tilde{X}_R \in B_r$, $f_{\mathcal{I}}(\tilde{X}_R, I)$ is positive definite and $\tilde{X}_R = I$ is the only critical point of f in B_r. Further, this implies that $V_1(\tilde{X}_R, \mathbf{x}_f)$ is positive definite in the set $\bar{B}_r = \{(\tilde{X}_R, \mathbf{x}_f) \in G \times \mathbb{R}^{n_f} \mid \ell(\tilde{X}_R, \mathbf{x}_f) \leq r\}$, where $\ell(\tilde{X}_R, \mathbf{x}_f) = d(\tilde{X}_R) + ||\mathbf{x}_f||_2$, and $\tilde{X}_R = I$ is the only critical point of f in $\bar{B}_r$.

A corollary to LaSalle's invariant set theorem will now be used to prove local asymptotic stability [94, p. 128]. Let $\mathcal{S} = \{(\tilde{X}_R, \mathbf{x}_f) \in \bar{B}_r \mid \dot{V}_1(\tilde{X}_R, \mathbf{x}_f) = 0\}$. It will now be shown that the only solution that can stay identically in $\mathcal{S}$ is the solution $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$. For all $(\tilde{X}_R, \mathbf{x}_f) \in \mathcal{S}$, $\mathbf{x}_f \equiv \mathbf{0}$. With $\mathbf{x}_f \equiv \mathbf{0}$, it follows that $\dot{\mathbf{x}}_f \equiv \mathbf{0}$. This implies that $\mathbf{B}_f \mathbf{e}_R \equiv \mathbf{0}$. Since $\mathbf{B}_f$ has full rank by assumption, $\mathbf{B}_f \mathbf{e}_R \equiv \mathbf{0}$ implies

that $\mathbf{e}_R \equiv \mathbf{0}$ and consequently $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) \equiv 0$. As the only critical point of f in $\bar{B}_r$ is the point $\tilde{X}_R = I$, it follows that $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) \equiv 0$ implies $\tilde{X}_R = I$. Thus, by Corollary 4.1 of [94, p. 128] the equilibrium point $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ is locally asymptotically stable. This proves item (i).

The proof of item (ii) continues as follows. By Corollary 6.29 of [17] there exists a constant $c > 0$ such that the set $\Omega_c = \{\tilde{X}_R \in G \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq c\}$ is compact and the only critical point of $f_{\mathcal{I}}$ in Ω_c is the point $\tilde{X}_R = I$. By Proposition 6.30 of [17], for all $c \in (0, L)$ there exists constants $0 < b_1 \leq b_2$ such that

$$b_1 ||\mathbf{e}_R||_2^2 \leq f_{\mathcal{I}}(\hat{X}, X) \leq b_2 ||\mathbf{e}_R||_2^2. \tag{6.26}$$

Consider again the function V_1 given in (6.25) and recall that $\dot{V}_1 \leq 0$. Therefore, $V_1(\tilde{X}_R(t), \mathbf{x}_f(t)) \leq V_1(\tilde{X}_R(0), \mathbf{x}_f(0))\ \forall t \geq 0$. Let $V_1(\tilde{X}_R(0), \mathbf{x}_f(0)) \leq c$, where $c \in (0, L)$. Then, $f_{\mathcal{I}}(\tilde{X}_R(t), I) \leq c\ \forall t \geq 0$ and $\tilde{X}_R \in \Omega_c\ \forall t \geq 0$. This implies that (6.26) is satisfied for all $t \geq 0$. Then, V_1 satisfies $\min\{b_1, \frac{1}{2}\underline{\lambda}(\mathbf{P}_f)\}||\mathbf{z}||_2^2 \leq V_1 \leq \max\{b_2, \frac{1}{2}\bar{\lambda}(\mathbf{P}_f)\}||\mathbf{z}||_2^2$, where $\underline{\lambda}(\cdot)$ and $\bar{\lambda}(\cdot)$ respectively denote the maximum and minimum eigenvalues of a matrix, and $\mathbf{z} = [\ ||\mathbf{e}_R||_2\ ||\mathbf{x}_f||_2\ ]^{\mathsf{T}}$. Thus, the fact that $V_1(\tilde{\mathbf{X}}(t), \mathbf{x}_f(t)) \leq c$ for all $t \geq 0$ implies that $\mathbf{e}_R$ and $\mathbf{x}_f$ are bounded for all $t \geq 0$.

Consider the Lyapunov function candidate

$$V_2(\tilde{X}_R, \mathbf{x}_f) = V_1(\tilde{X}_R, \mathbf{x}_f) - a\mathbf{e}_R^{\mathsf{T}}\mathbf{C}_f\mathbf{x}_f, \tag{6.27}$$

where $a \in (0, \infty)$. Let $\gamma = ||\mathbf{C}_f||_{\mathsf{F}}$, then V_2 satisfies $\mathbf{z}^{\mathsf{T}}\mathbf{W}_1\mathbf{z} \leq V_2 \leq \mathbf{z}^{\mathsf{T}}\mathbf{W}_2\mathbf{z}$, where

$$\mathbf{W}_1 = \begin{bmatrix} b_1 & -\frac{1}{2}a\gamma \\ -\frac{1}{2}a\gamma & \frac{1}{2}\underline{\lambda}(\mathbf{P}_f) \end{bmatrix}, \text{ and } \mathbf{W}_2 = \begin{bmatrix} b_2 & \frac{1}{2}a\gamma \\ \frac{1}{2}a\gamma & \frac{1}{2}\bar{\lambda}(\mathbf{P}_f) \end{bmatrix}.$$

The derivative with respect to time of V_2 satisfies

$$\begin{aligned} \dot{V}_2 &= \dot{V}_1 - a\mathbf{x}_f^{\mathsf{T}}\mathbf{C}_f^{\mathsf{T}}\dot{\mathbf{e}}_R - a\mathbf{e}^{\mathsf{T}}\mathbf{C}_f\dot{\mathbf{x}}_f \\ &\leq -\tfrac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f - a\mathbf{x}_f^{\mathsf{T}}\mathbf{C}_f^{\mathsf{T}}\dot{\mathbf{e}}_R - a\mathbf{e}_R^{\mathsf{T}}\mathbf{C}_f\dot{\mathbf{x}}_f. \end{aligned}$$

From Corollary 6.7, the derivative with respect to time of $\mathbf{e}_R$ is given by $\dot{\mathbf{e}}_R = -\mathbf{H}(\tilde{X}_R)\mathbf{u}_R - \boldsymbol{\xi}$, where $S(\boldsymbol{\xi}) = Q(S(\mathbf{u}_R), S(\mathbf{e}_R))$ and $||\boldsymbol{\xi}||_2 \leq m||\mathbf{u}_R||_2||\mathbf{e}_R||_2$. Given that $\mathbf{x}_f$ and $\mathbf{e}_R$ are bounded, it follows that $\mathbf{u}_R$ is bounded as well. As Ω_c is compact it follows that the norm of $\mathbf{H}(\tilde{X}_R)$ is bounded [17]. Define $m_1 = \sup_{\tilde{X}_R \in \Omega_c} ||\mathbf{H}(\tilde{X}_R)||_{\mathsf{F}}$,

$m_2 = m \cdot \sup ||\mathbf{u}_R||_2$, $\beta = ||\mathbf{D}_f||_\mathsf{F}$, and $\epsilon = \underline{\lambda}(\mathbf{Q}_f)$. Then it can be shown that

$$\begin{aligned}\dot{V}_2 \;\leq\; & -\tfrac{1}{2}\epsilon||\mathbf{x}_f||^2 + am_1\gamma^2||\mathbf{x}_f||_2^2 + am_1\beta\gamma||\mathbf{x}_f||_2||\mathbf{e}_R||_2 \\ & + a\gamma m_2||\mathbf{x}_f||_2||\mathbf{e}_R||_2 - a\mathbf{e}^\mathsf{T}\mathbf{C}_f\mathbf{A}_f\mathbf{x}_f \\ & - a\mathbf{e}_R^\mathsf{T}\mathbf{C}_f\mathbf{B}_f\mathbf{e}_R.\end{aligned}$$

Recall from (6.20b) $\mathbf{C}_f = \mathbf{B}_f^\mathsf{T}\mathbf{P}_f$. Let $\boldsymbol{\Upsilon} = \mathbf{B}_f^\mathsf{T}\mathbf{P}_f\mathbf{B}_f$ and note that since $\mathbf{B}_f$ is assumed full rank and $\mathbf{P}_f > 0$ it follows that $\boldsymbol{\Upsilon} > 0$. Define $\delta = ||\mathbf{A}_f||_\mathsf{F}$, then

$$\begin{aligned}\dot{V}_2 \leq & -(\tfrac{1}{2}\epsilon - am_1\gamma^2)||\mathbf{x}_f||_2^2 - a\underline{\lambda}(\boldsymbol{\Upsilon})||\mathbf{e}_R||_2^2 \\ & + a\gamma(m_2 + m_1\beta + \delta)||\mathbf{x}_f||_2||\mathbf{e}_R||_2. \end{aligned} \tag{6.28}$$

Equation (6.28) is equivalent to $\dot{V}_2 \leq -\mathbf{z}^\mathsf{T}\mathbf{W}_3\mathbf{z}$, where

$$\mathbf{W}_3 = \begin{bmatrix} a\underline{\lambda}(\boldsymbol{\Upsilon}) & -a\kappa \\ -a\kappa & \frac{1}{2}\epsilon - am_1\gamma^2 \end{bmatrix},$$

and $\kappa = \frac{1}{2}a\gamma(m_2 + m_1\beta + \delta)$. It can be shown that, provided

$$a < \min\left\{ \sqrt{\frac{2b_1\underline{\lambda}(\mathbf{P}_f)}{\gamma^2}}, \frac{\frac{1}{2}\epsilon\underline{\lambda}(\boldsymbol{\Upsilon})}{\underline{\lambda}(\boldsymbol{\Upsilon})m_1\gamma^2 + \kappa^2} \right\},$$

the matrices $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{W}_3$ are positive definite. Therefore,

$$\begin{aligned}\underline{\lambda}(\mathbf{d}_1)||\mathbf{z}||_2^2 \leq V_2 \leq \bar{\lambda}(\mathbf{d}_2)||\mathbf{z}||_2^2 \\ \dot{V}_2 \leq -\underline{\lambda}(\mathbf{d}_3)||\mathbf{z}||_2^2,\end{aligned}$$

which implies that

$$||\mathbf{z}(t)||_2 \leq \left(\frac{\underline{\lambda}(\mathbf{d}_1)}{\bar{\lambda}(\mathbf{d}_2)}\right)^{\frac{1}{2}} ||\mathbf{z}(0)||_2 \exp\left(-\frac{\underline{\lambda}(\mathbf{d}_3)}{2\bar{\lambda}(\mathbf{d}_2)}t\right).$$

Therefore, trajectories of $(||\mathbf{e}_R||_2, ||\mathbf{x}_f||_2)$ exponentially approach $(0,0)$. Due to the fact that trajectories of $\tilde{X}_R$ remain in Ω_c for all $t \geq 0$, $\mathbf{e}_R \to \mathbf{0}$ implies that $\tilde{X}_R \to I$ as $t \to \infty$. $\square$

6.4.2 Stability of the Left Observer

Theorem 6.9. *Consider trajectories of* $(\tilde{X}_R, \mathbf{x}_f)$ *under* (6.19). *Let* $f_{\mathcal{I}}$ *be a bi-invariant error function about* X *and take a bi-invariant Riemannian metric. Let* $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f)$ *be SPR where* $\mathbf{B}_f$ *has full rank, and let* $\mathbf{D}_f \geq 0$. *Assume the following:*

(A1) there is no noise associated with the group velocity measurement or the partial state measurements such that $d = 0$ *and* $N = \mathcal{I}$;

(A2) there exists a faithful representation of the Lie group G *as a matrix Lie group;*

(A3) the matrix representation of the state X *remains bounded with respect to* $||\cdot||_{\mathsf{F}}$ *for all* $t \geq 0$ *and* v *remains bounded with respect to* $||\cdot||_{\mathfrak{g}}$ *for all* $t \geq 0$.

Define $L = \sup\{c \in \mathbb{R} \mid \tilde{X}_R \in \Omega_c \setminus \{I\} \implies \mathbf{e}_L \neq \mathbf{0}\}$, *where* $\Omega_c = \{\tilde{X}_R \in G \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq c\}$. *Then the equilibrium point* $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ *is locally uniformly asymptotically stable.*

Proof. First, item (i) will be proved. Consider the Lyapunov function candidate

$$V_1(\tilde{X}_R, \mathbf{x}_f) = f_{\mathcal{I}}(\tilde{X}_R, I) + \mathcal{L}(\mathbf{x}_f). \tag{6.29}$$

The derivative with respect to time of V_1 is

$$\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \;=\; \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f).$$

From Proposition 6.2, $\dot{\tilde{X}}_R = -\mathrm{Ad}_{\hat{X}}(u_L)\tilde{X}_L$ and $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) = \mathrm{Ad}_{\hat{X}}(e_L)\tilde{X}_R$. Therefore,

$$\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \;=\; -\langle \mathrm{Ad}_{\hat{X}}(e_L)\tilde{X}_L, \mathrm{Ad}_{\hat{X}}(u_L)\tilde{X}_L \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f).$$

Due to the bi-invariance property of the Riemannian metric,

$$\langle \mathrm{Ad}_{\hat{X}}(e_L)\tilde{X}_L, \mathrm{Ad}_{\hat{X}}(u_L)\tilde{X}_L \rangle_{\tilde{X}_R} = \langle e_L, u_L \rangle, \tag{6.30}$$

and by application of Corollary 6.6, $\dot{V}_1$ is given by

$$\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) = -\mathbf{e}_L^{\mathsf{T}}\mathbf{u}_L - \tfrac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f + \mathbf{e}_L^{\mathsf{T}}\mathbf{u}_L - \mathbf{e}_L^{\mathsf{T}}\mathbf{D}_f\mathbf{e}_L.$$

Consequently, $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) = -\frac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f - \mathbf{e}_L^{\mathsf{T}}\mathbf{D}_f\mathbf{e}_L$. By assumption $\mathbf{D}_f \geq 0$, which implies $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \leq -\frac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f$ and thus $\dot{V}_1(\tilde{X}_R, \mathbf{x}_f) \leq 0$ and $V_1(\tilde{X}_R(t), \mathbf{x}_f(t)) \leq$

$V_1(\tilde{X}_R(0), \mathbf{x}_f(0))$ for all $t \geq 0$. By Remark 11.11 of [17], $f_{\mathcal{I}}$ is locally positive definite and $\tilde{X} = I$ is an isolated critical point of $f_{\mathcal{I}}$. By assumption, there exists a faithful representation of G as a matrix Lie group. This implies that there exists $m > 0$ and a mapping $\mathbf{\Phi} : G \to GL(m)$ such that $\mathbf{\Phi}(G)$ is a matrix Lie group [16]. Following the proof of Theorem 5.1 in [16], this implies that there exists a set $B_r = \{\tilde{X}_R \in G \mid d(\tilde{X}_R) \leq r\}$ about $\tilde{X}_R = I$, where $d(\tilde{X}_R) = ||\mathbf{1} - \mathbf{\Phi}(\tilde{X}_R)||_{\mathsf{F}}$, such that for all $\tilde{X}_R \in B_r$, $f_{\mathcal{I}}(\tilde{X}_R, I)$ is positive definite and $\tilde{X}_R = I$ is the only critical point of f in B_r. Further, this implies that $V_1(\tilde{X}_R, \mathbf{x}_f)$ is positive definite in the set $\bar{B}_r = \{(\tilde{X}_R, \mathbf{x}_f) \in G \times \mathbb{R}^{n_f} \mid \ell(\tilde{X}_R, \mathbf{x}_f) \leq r\}$, where $\ell(\tilde{X}_R, \mathbf{x}_f) = d(\tilde{X}_R) + ||\mathbf{x}_f||_2$, and $\tilde{X}_R = I$ is the only critical point of f in $\bar{B}_r$. Consequently, for all $t \geq 0$ and all $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \bar{B}_r$ the Lyapunov function V_1 is positive definite and $\dot{V}_1 \leq 0$ and by Theorem 4.8 of [94] the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is uniformly stable.

Again, following the proof of Theorem 5.1 in [16], choose

$$\alpha < \min_{\ell(\tilde{X}_R, \mathbf{x}_f) = r} V_1(\tilde{X}_R, \mathbf{x}_f),$$

and define $\mathscr{L}_\alpha = \{(\tilde{X}_R, \mathbf{x}_f) \in \bar{B}_r \mid V_1(\tilde{X}_R, \mathbf{x}_f) \leq \alpha\}$. Then, $\mathscr{L}_\alpha \subset \bar{B}_r$ and all trajectories of $(\tilde{X}_R, \mathbf{x}_f)$ starting in $\mathscr{L}_\alpha$ remain in $\mathscr{L}_\alpha$ for all $t \geq 0$ [94]. Consequently, trajectories of $(\tilde{X}_R, \mathbf{x}_f)$ remain bounded with respect to $\ell(\tilde{X}_R, \mathbf{x}_f)$, which implies $\tilde{X}_R$ remains bounded with respect to $d(\cdot)$ and $\mathbf{x}_f$ remains bounded with respect to $||\cdot||_2$.

As $(\tilde{X}_R, \mathbf{x}_f) \in \mathscr{L}_\alpha$ for all $t \geq 0$, it follows that $\tilde{X}_R \in \Omega_\alpha$ for all $t \geq 0$, where $\Omega_\alpha = \{X \in B_r \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq \alpha\}$. Moreover, $\mathscr{L}_\alpha \subset \bar{B}_r$ implies that the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is $\tilde{X}_R = I$ and therefore the only critical point in Ω_α is $\tilde{X} = I$. Consequently, $\Omega_\alpha \subset \Omega_L$ and therefore by Proposition 6.30 of [17] there exists constants b_1 and b_2 such that (6.26) is satisfied. This implies that $\mathbf{e}_R$ is bounded for all $t \geq 0$.

Taking the second derivative with respect to time of V_1 yields $\ddot{V}_1 = -\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\dot{\mathbf{x}}_f - 2\mathbf{e}_L^{\mathsf{T}}\mathbf{D}_f\dot{\mathbf{e}}_L$. By Corollary 6.7, $\dot{\mathbf{e}}_R = \mathbf{H}(\tilde{X}_R)\mathbf{q} - \boldsymbol{\xi}$ where $S(\mathbf{q}) = -\mathrm{Ad}_{\hat{X}}(u_L)$ and $S(\boldsymbol{\xi}) = Q(-\mathrm{Ad}_{\hat{X}}(u_L), e_R)$. By assumption $||\mathbf{\Phi}(X)||_{\mathsf{F}}$ is bounded, which implies that $||\mathbf{\Phi}(\hat{X})||_{\mathsf{F}}$ is bounded due to the fact that $\tilde{X}_R$ is bounded. Given that $\tilde{X}_R$, $\hat{X}$, $\mathbf{x}_f$, and $\mathbf{e}_R$ are bounded it follows that $\mathbf{e}_L$, $\mathbf{u}_f$, $\dot{\mathbf{x}}_f$ and $\dot{\mathbf{e}}_R$ are bounded. Recall, $e_L = \mathrm{Ad}_{\hat{X}^{-1}}(e_R)$ and therefore $\dot{e}_L = \mathrm{Ad}_{\hat{X}^{-1}}(\dot{e}_R) + [\mathrm{Ad}_{\hat{X}^{-1}}(e_R), v - u_L]$ [17, p. 307]. As v is bounded by assumption, it follows that $\dot{e}_L$ is bounded and thus $\dot{\mathbf{e}}_L$ is bounded as well. Thus, $\ddot{V}_1$ is bounded and therefore $\dot{V}_1$ is uniformly continuous. By application of Barbalat's Lemma $\dot{V}_1 \to 0$ as $t \to \infty$ and thus $\mathbf{x}_f \to \mathbf{0}$ as $t \to \infty$. Taking the derivative with respect to time of $\dot{\mathbf{x}}_f$ gives $\ddot{\mathbf{x}}_f = \mathbf{A}_f^2\mathbf{x}_f + \mathbf{A}_f\mathbf{B}_f\mathbf{e}_L + \mathbf{B}_f\dot{\mathbf{e}}_L$, which is bounded. Applying Barbalat's Lemma, $\dot{\mathbf{x}}_f \to \mathbf{0}$. This fact, along with the assumption that $\mathbf{B}_f$ has full

rank implies that $\mathbf{e}_L \to \mathbf{0}$ and $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) \to 0$. Since the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is the point $\tilde{X}_R = I$, it follows that $\tilde{X}_R \to I$ as $t \to \infty$.

It has been shown that for all $(\tilde{X}_R(0), \mathbf{x}_f(0)) \in \mathscr{L}_\alpha$, $(\tilde{X}_R, \mathbf{x}_f) \to (\mathbf{1}, \mathbf{0})$ as $t \to \infty$. As V_1 is positive definite on $\mathscr{L}_\alpha$ there exists a class $\mathcal{K}$ function ϕ such that $\phi(\ell(\tilde{X}_R, \mathbf{x}_f)) \leq V(\tilde{X}_R, \mathbf{x}_f)$ for all $(\tilde{X}_R, \mathbf{x}_f) \in \mathscr{L}_\alpha$ [94]. Thus, $(\tilde{X}_R, \mathbf{x}_f) \to (I, \mathbf{0})$ as $t \to \infty$ for all $\ell(\tilde{X}_R(0), \mathbf{x}_f(0)) < \phi^{-1}(\alpha)$ and therefore the equilibrium point $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ is uniformly convergent [95]. This, along with the fact that the equilibrium point is uniformly stable, shows that $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ is locally uniformly asymptotically stable. $\square$

6.4.3 Discussion of Stability Results

The stability results presented in Theorems 6.8 and 6.9 are valid for all $\mathbf{H}(s) = \mathbf{H}_{\text{spr}}(s) + \mathbf{D}_f$, where $\mathbf{H}_{\text{spr}}(s)$ is nonzero and $\mathbf{D}_f \geq 0$. When $\mathbf{H}_{\text{spr}}(s) = \mathbf{0}$ a similar result can be found by further restricting $\mathbf{D}_f$ to the set of positive definite matrices, that is $\mathbf{D}_f > 0$. This can be accomplished by taking $f_{\mathcal{I}}(\tilde{X}, I)$ as the Lyapunov function and performing a similar analysis. This result is not presented here as it is equivalent to the stability results presented previously in [14].

The restriction on the set of initial conditions in item (ii) of Theorem 6.8 can be interpreted as an estimate of the region of attraction of the equilibrium point $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$. Specifically, the estimate of the region of attraction is $\{(\tilde{X}_R, \mathbf{x}_f) \in G \times \mathbb{R}^{n_f} \mid V_1(\tilde{X}_R, \mathbf{x}_f) \leq c\}$, where $c < L$. It also follows, by Remark 6.13 of [17], that if $\tilde{X}_R = I$ is the only critical point of f, then the equilibrium point $(\tilde{X}_R, \mathbf{x}_f) = (I, \mathbf{0})$ is globally asymptotically stable. Often it is the case that f will have multiple critical points. In these instances, it is not possible to demonstrate global asymptotic stability. However, it may be possible to demonstrate almost global stability by placing further restrictions on the the error function as is done in [14].

6.5 Disturbance Observer

In the proofs in the previous section, it was assumed that the group velocity v is measured exactly. Suppose that the noise associated with v_y in (6.3a), $d = S(\mathbf{d})$ where $\mathbf{d} \in \mathbb{R}^n$, is composed of a linear combination of constant and harmonic signals. As such, $\mathbf{d}$ may be written as the output of the linear system

$$\dot{\mathbf{x}}_d = \mathbf{A}_d \mathbf{x}_d, \quad \mathbf{d} = \mathbf{C}_d \mathbf{x}_d, \tag{6.31}$$

where $\mathbf{x}_d \in \mathbb{R}^{n_d}$ and $\mathbf{A}_d$ is skew-symmetric [96].

6.5.1 Right Observer with Disturbance Rejection

Let $\hat{\mathbf{x}}_d$ denote the estimate of $\mathbf{x}_d$ and consider the right observer with disturbance rejection given by

$$\begin{aligned}
\dot{\hat{X}} &= \hat{X}v_y - \hat{X}\hat{d} - u_R\hat{X}, && \text{(6.32a)}\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && \text{(6.32b)}\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && \text{(6.32c)}\\
\dot{\hat{\mathbf{x}}}_d &= \mathbf{A}_d\hat{\mathbf{x}}_d + \rho\mathbf{C}_d^{\mathsf{T}}\bar{\mathbf{e}}_R, && \text{(6.32d)}\\
\hat{\mathbf{d}} &= \mathbf{C}_d\hat{\mathbf{x}}_d, && \text{(6.32e)}
\end{aligned}$$

where $\rho > 0$, and $\hat{d} = S(\hat{\mathbf{d}}) \in \mathfrak{g}$. The input to (6.32d), $\bar{\mathbf{e}}_R \in \mathbb{R}^n$, is such that $S(\bar{\mathbf{e}}_R) = \mathrm{Ad}^*_{\hat{X}}(S(\mathbf{e}_R))$ where $\mathrm{Ad}^*_{\hat{X}}(\cdot)$ is the adjoint of the linear map $\mathrm{Ad}_{\hat{X}}(\cdot)$ such that $\langle u, \mathrm{Ad}_X(v)\rangle = \langle \mathrm{Ad}^*_X(u), v\rangle$ for all $v, u \in \mathfrak{g}$ and $X \in G$. The Lie group observer presented in (6.32) is similar to the observer presented in (6.11). However, the estimate of the disturbance, given by the disturbance observer in (6.32d) and (6.32e), is subtracted from the velocity measurement. When the disturbance is a constant bias the proposed method is equivalent to the bias estimations methods previously considered in Sec. 4.4.2, Sec. 5.5, and in [15]. Similar disturbance observers are used in [97] in the context of attitude estimation as well as in [96] in the context of spacecraft attitude control.

The error dynamics associated with (6.32) will be needed in the stability analysis that follows. The error associated with the disturbance estimate is defined as $\tilde{\mathbf{x}}_d = \mathbf{x}_d - \hat{\mathbf{x}}_d$ and $\tilde{d} = S(\tilde{\mathbf{d}})$ where $\tilde{\mathbf{d}} = \mathbf{d} - \hat{\mathbf{d}} \in \mathbb{R}^n$.

Proposition 6.10. *Consider trajectories of* $(\hat{X}, \mathbf{x}_f, \mathbf{x}_d)$ *under* (6.32). *Let* $f_N : G \times G \to \mathbb{R}$ *be a right-invariant cost function and take a right-invariant Riemannian metric on* G. *Assume that group velocity measurements and partial state measurements are available as in* (6.3) *and assume that* $d = S(\mathbf{d})$ *is given by* (6.31). *Then, the*

dynamics associated with $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ *are nonautonomous and are given by*

$$\begin{aligned}
\dot{\tilde{X}}_R &= \tilde{X}_R \mathrm{Ad}_X(\tilde{d}) - u_R \tilde{X}_R, && \text{(6.33a)}\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f \mathbf{x}_f + \mathbf{B}_f \mathbf{e}_R, && \text{(6.33b)}\\
\mathbf{u}_R &= \mathbf{C}_f \mathbf{x}_f + \mathbf{D}_f \mathbf{e}_R, && \text{(6.33c)}\\
\dot{\tilde{\mathbf{x}}}_d &= \mathbf{A}_d \tilde{\mathbf{x}}_d - \rho \mathbf{C}_d^{\mathsf{T}} \bar{\mathbf{e}}_R, && \text{(6.33d)}\\
\tilde{\mathbf{d}} &= \mathbf{C}_d \tilde{\mathbf{x}}_d. && \text{(6.33e)}
\end{aligned}$$

Proof By Lemma 3.1 the time derivative of $\tilde{X}_R$ satisfies $\dot{\tilde{X}}_R = T_{\hat{X}} R_{X^{-1}}(\hat{X}d - \hat{X}\hat{d} + u_R \hat{X})$. It follows then that

$$\begin{aligned}
\dot{\tilde{X}}_R &= T_{\hat{X}} R_{X^{-1}}(\hat{X}\tilde{d} - u_R \hat{X})\\
&= T_{\hat{X}} R_{X^{-1}} T_I L_{\hat{X}} \tilde{d} - T_{\hat{X}} R_{X^{-1}} u_R \hat{X}\\
&= T_{X^{-1}} L_{\hat{X}} T_I R_{X^{-1}} \tilde{d} - u_R \tilde{X}_R\\
&= T_I L_{\tilde{X}_R} T_{X^{-1}} L_X T_I R_{X^{-1}} \tilde{d} - u_R \tilde{X}_R\\
&= T_I L_{\tilde{X}_R} \mathrm{Ad}_X(\tilde{d}) - u_R \tilde{X}_R\\
&= \tilde{X}_R \mathrm{Ad}_X(\tilde{d}) - u_R \tilde{X}_R.
\end{aligned}$$

Taking the time derivative of $\tilde{\mathbf{x}}_d$ gives

$$\begin{aligned}
\dot{\tilde{\mathbf{x}}}_d &= \dot{\mathbf{x}}_d - \dot{\hat{\mathbf{x}}}_d\\
&= \mathbf{A}_d \mathbf{x}_d - \mathbf{A}_d \hat{\mathbf{x}}_d - \rho \mathbf{C}_d^{\mathsf{T}} \bar{\mathbf{e}}_R\\
&= \mathbf{A}_d \tilde{\mathbf{x}}_d - \rho \mathbf{C}_d^{\mathsf{T}} \bar{\mathbf{e}}_R.
\end{aligned}$$

The error in the disturbance estimate is $\tilde{\mathbf{d}} = \mathbf{d} - \hat{\mathbf{d}} = \mathbf{C}_d \mathbf{x}_d - \mathbf{C}_d \hat{\mathbf{x}}_d = \mathbf{C}_d \tilde{\mathbf{x}}_d$. The dynamics are nonautonomous due to the presence of the time dependent variable X in the expression for $\dot{\tilde{X}}_R$. □

6.5.2 Left Observer with Disturbance Rejection

The left observer with disturbance rejection takes the form

$$\begin{aligned}
\dot{\hat{X}} &= \hat{X}v_y - \hat{X}\hat{d} - \hat{X}u_L, && (6.34a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && (6.34b)\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && (6.34c)\\
\dot{\hat{\mathbf{x}}}_d &= \mathbf{A}_d\hat{\mathbf{x}}_d + \rho\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L, && (6.34d)\\
\hat{\mathbf{d}} &= \mathbf{C}_d\hat{\mathbf{x}}_d, && (6.34e)
\end{aligned}$$

where again $\rho > 0$, and $\hat{d} = S(\hat{\mathbf{d}}) \in \mathfrak{g}$.

Proposition 6.11. *Consider trajectories of* $(\hat{X}, \mathbf{x}_f, \mathbf{x}_d)$ *under* (6.34). *Let* $f_N : G \times G \to \mathbb{R}$ *be a bi-invariant cost function and take a bi-invariant Riemannian metric on* G. *Assume that group velocity measurements and partial state measurements are available as in* (6.3) *and assume that* $d = S(\mathbf{d})$ *is given by* (6.31). *Then, the dynamics associated with* $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ *are nonautonomous and are given by*

$$\begin{aligned}
\dot{\tilde{X}}_L &= \tilde{X}_L \mathrm{Ad}_X(\tilde{d} - u_L), && (6.35a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && (6.35b)\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && (6.35c)\\
\dot{\tilde{\mathbf{x}}}_d &= \mathbf{A}_d\tilde{\mathbf{x}}_d - \rho\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L, && (6.35d)\\
\tilde{\mathbf{d}} &= \mathbf{C}_d\tilde{\mathbf{x}}_d. && (6.35e)
\end{aligned}$$

Proof By Lemma 3.1 the time derivative of $\tilde{X}_R$ satisfies $\dot{\tilde{X}}_R = T_{\hat{X}}R_{X^{-1}}(\hat{X}d - \hat{X}\hat{d} + \hat{X}u_L)$. It follows then that

$$\begin{aligned}
\dot{\tilde{X}}_R &= T_{\hat{X}}R_{X^{-1}}(\hat{X}(\tilde{d} - u_L))\\
&= T_{\hat{X}}R_{X^{-1}}T_IL_{\hat{X}}(\tilde{d} - u_L)\\
&= T_{X^{-1}}L_{\hat{X}}T_IR_{X^{-1}}(\tilde{d} - u_L)\\
&= T_IL_{\tilde{X}_R}T_{X^{-1}}L_XT_IR_{X^{-1}}(\tilde{d} - u_L)\\
&= T_IL_{\tilde{X}_R}\mathrm{Ad}_X(\tilde{d} - u_L)\\
&= \tilde{X}_R\mathrm{Ad}_X(\tilde{d} - u_L).
\end{aligned}$$

Taking the time derivative of $\tilde{\mathbf{x}}_d$ gives

$$\begin{aligned}\dot{\tilde{\mathbf{x}}}_d &= \dot{\mathbf{x}}_d - \dot{\hat{\mathbf{x}}}_d \\ &= \mathbf{A}_d\mathbf{x}_d - \mathbf{A}_d\hat{\mathbf{x}}_d - \rho\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L \\ &= \mathbf{A}_d\tilde{\mathbf{x}}_d - \rho\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L.\end{aligned}$$

The error in the disturbance estimate is $\tilde{\mathbf{d}} = \mathbf{d} - \hat{\mathbf{d}} = \mathbf{C}_d\mathbf{x}_d - \mathbf{C}_d\hat{\mathbf{x}}_d = \mathbf{C}_d\tilde{\mathbf{x}}_d$. The dynamics are nonautonomous due to the presence of the time dependent variable X in the expression for $\dot{\tilde{X}}_R$. $\square$

6.5.3 Stability Results

The stability of the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ for the right and left observers with disturbance rejection is established in Theorems 6.12 and 6.13, respectively. As before, it is required that f_N be an error function and that $\mathbf{H}(s) = \mathbf{H}_{\text{spr}}(s) + \mathbf{D}_f$, where $\mathbf{H}_{\text{spr}}(s)$ is SPR and $\mathbf{D}_f \geq 0$.

6.5.3.1 Stability of the Right Observer with Disturbance Rejection

Theorem 6.12. *Consider trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ under (6.33). Let f_N be a right invariant error function about X and take a right-invariant Riemannian metric. Let $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f)$ be SPR, where $\mathbf{B}_f$ has full rank and $\mathbf{D}_f \geq 0$. Assume the following:*

(A1) there is no noise associated with the group velocity measurement or the partial state measurements such that $d = 0$ and $N = \mathcal{I}$;

(A2) there exists a faithful representation of the Lie group G as a matrix Lie group;

(A3) $\mathbf{\Phi}(X)$ and v are bounded with respect to $||\cdot||_{\mathsf{F}}$ and $||\cdot||_{\mathfrak{g}}$, respectively;

(A4) $\mathbf{C}_d$ has full rank.

Define $L = \sup\{c \in \mathbb{R} \mid \tilde{X}_R \in \Omega_c \setminus \{I\} \implies \mathbf{e}_R \neq \mathbf{0}\}$, where $\Omega_c = \{\tilde{X}_R \in G \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq c\}$. Then the following statements hold:

(i) the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is locally uniformly asymptotically stable;

(ii) trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ converge asymptotically to $(I, \mathbf{0}, \mathbf{0})$ for all initial conditions satisfying $V_3(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) < L$, where V_3 is defined in (6.36).

Proof. The proof of item (i) in Theorem 6.12 follows in a similar manner to the proof of Theorem 5.1 in [16]. Consider the Lyapunov function candidate

$$V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = f_{\mathcal{I}}(\tilde{X}_R, I) + \mathcal{L}(\mathbf{x}_f) + \tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d. \tag{6.36}$$

The derivative with respect to time of the third term, $\frac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d$, is

$$\frac{\mathrm{d}}{\mathrm{d}t}\tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d = \tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}(\mathbf{A}_d^{\mathsf{T}} + \mathbf{A}_d)\tilde{\mathbf{x}}_d - \tilde{\mathbf{x}}_d^{\mathsf{T}}\mathbf{C}_d^{\mathsf{T}}\bar{\mathbf{e}}_R.$$

Recall that $\mathbf{A}_d$ is skew-symmetric and therefore $\mathbf{A}_d^{\mathsf{T}} + \mathbf{A}_d = \mathbf{0}$. Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d = -\tilde{\mathbf{x}}_d^{\mathsf{T}}\mathbf{C}_d^{\mathsf{T}}\bar{\mathbf{e}}_R = -\tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R.$$

The derivative with respect to time of V_3 can therefore be written as

$$\dot{V}_3 = \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R. \tag{6.37}$$

Substituting (6.33a) into (6.37) yields

$$\begin{aligned}
\dot{V}_3 &= \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= \langle S(\mathbf{e}_R)\tilde{X}_R, \tilde{X}_R \mathrm{Ad}_X(\tilde{d}) \rangle_{\tilde{X}_R} - \langle S(\mathbf{e}_R)\tilde{X}_R, S(\mathbf{u}_R)\tilde{X}_R \rangle_{\tilde{X}_R} \\
&\quad + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= -\langle S(\mathbf{e}_R), S(\mathbf{u}_R) \rangle + \langle S(\mathbf{e}_R), \mathrm{Ad}_{\hat{X}}(\tilde{d}) \rangle \\
&\quad + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= -\mathbf{e}_R^{\mathsf{T}}\mathbf{u}_R + \langle \mathrm{Ad}^*_{\hat{X}}(S(\mathbf{e}_R)), S(\tilde{\mathbf{d}}) \rangle + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= -\mathbf{e}_R^{\mathsf{T}}\mathbf{u}_R + \langle S(\bar{\mathbf{e}}_R), S(\tilde{\mathbf{d}}) \rangle + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= -\mathbf{e}_R^{\mathsf{T}}\mathbf{u}_R + \bar{\mathbf{e}}_R^{\mathsf{T}}\tilde{\mathbf{d}} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\bar{\mathbf{e}}_R \\
&= -\mathbf{e}_R^{\mathsf{T}}\mathbf{u}_R + \dot{\mathcal{L}}(\mathbf{x}_f).
\end{aligned}$$

By Corollary 6.6, $\dot{V}_3 = -\frac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f - \mathbf{e}_R^{\mathsf{T}}\mathbf{D}_f\mathbf{e}_R$. Therefore, $\dot{V}_3 \leq 0$. As $f_{\mathcal{I}}$ is an error function it follows that $f_{\mathcal{I}}$ is locally positive definite and $\tilde{X}_R = I$ is an isolated critical point of $f_{\mathcal{I}}$ [17]. By assumption, there exists a faithful representation of G as a matrix Lie group. This implies that there exists $m > 0$ and a mapping $\mathbf{\Phi} : G \to GL(m)$ such that $\mathbf{\Phi}(G)$ is a matrix Lie group [16]. Following the proof of Theorem 5.1 in [16], this implies that there exists a set $B_r = \{\tilde{X}_R \in G \mid d(\tilde{X}_R) \leq r\}$ about $\tilde{X}_R = I$, where $d(\tilde{X}_R) = ||\mathbf{1} - \mathbf{\Phi}(\tilde{X}_R)||_{\mathsf{F}}$, such that for all $\tilde{X}_R \in B_r$, $f_{\mathcal{I}}(\tilde{X}_R, I)$ is positive definite and

$\tilde{X}_R = I$ is the only critical point of $f_{\mathcal{I}}$ in B_r. Further, this implies that $V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}})$ is positive definite in the set $\bar{B}_r = \{(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in G \times \mathbb{R}^{n_f} \times \mathbb{R}^{n_d} \mid \ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \leq r\}$, where $\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = d(\tilde{X}_R) + ||\mathbf{x}_f||_2 + ||\tilde{\mathbf{x}}_d||_2$, and $\tilde{X}_R = I$ is the only critical point of $f_{\mathcal{I}}$ in the set $\bar{B}_r$. Consequently, for all $t \geq 0$ and all $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \bar{B}_r$ the Lyapunov function V_3 is positive definite and $\dot{V}_3 \leq 0$ and by Theorem 4.8 of [94] the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is uniformly stable.

Again, following the proof of Theorem 5.1 in [16], choose

$$\alpha < \min_{\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = r} V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d),$$

and define $\mathscr{L}_\alpha = \{(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \bar{B}_r \mid V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \leq \alpha\}$. Then, $\mathscr{L}_\alpha \subset \bar{B}_r$ and all trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ starting in $\mathscr{L}_\alpha$ remain in $\mathscr{L}_\alpha$ for all $t \geq 0$ [94]. Consequently, trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ remain bounded with respect to $\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$, which implies $\tilde{X}_R$ remains bounded with respect to $d(\cdot)$ and $\mathbf{x}_f, \tilde{\mathbf{x}}_d$ remain bounded with respect to $|| \cdot ||_2$.

As $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \mathscr{L}_\alpha$ for all $t \geq 0$, it follows that $\tilde{X}_R \in \Omega_\alpha$ for all $t \geq 0$, where $\Omega_\alpha = \{X \in B_r \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq \alpha\}$. Moreover, $\mathscr{L}_\alpha \subset \bar{B}_r$ implies that the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is $\tilde{X}_R = I$ and therefore the only critical point in Ω_α is $\tilde{X}_R = I$. Consequently, $\Omega_\alpha \subset \Omega_L$ and therefore by Proposition 6.30 of [17] there exists constants b_1 and b_2 such that (6.26) is satisfied. This implies that $\mathbf{e}_R$ is bounded for all $t \geq 0$.

Taking the second derivative with respect to time of V_3 yields $\ddot{V}_3 = -\mathbf{x}_f^\mathsf{T} \mathbf{Q}_f \dot{\mathbf{x}}_f - 2\mathbf{e}_R^\mathsf{T} \mathbf{D}_f \dot{\mathbf{e}}_R$. By Corollary 6.7, $\dot{\mathbf{e}}_R = \mathbf{H}(\tilde{X}_R)\mathbf{q} - \boldsymbol{\xi}$ where $S(\mathbf{q}) = \mathrm{Ad}_{\tilde{X}}(S(\tilde{\mathbf{d}})) - S(\mathbf{u}_R)$ and $S(\boldsymbol{\xi}) = Q(S(\mathbf{q}), S(\mathbf{e}_R))$. Given that $\tilde{X}_R$, $\mathbf{x}_f$, $\tilde{\mathbf{x}}_d$, and $\mathbf{e}_R$ are bounded and X is bounded by assumption, it follows that $\dot{\mathbf{x}}_f$ and $\dot{\mathbf{e}}_R$ are bounded. Thus, $\ddot{V}_3$ is bounded and therefore $\dot{V}_3$ is uniformly continuous. By application of Barbalat's Lemma $\dot{V}_3 \to 0$ as $t \to \infty$ and thus $\mathbf{x}_f \to \mathbf{0}$ as $t \to \infty$. Taking the derivative with respect to time of $\dot{\mathbf{x}}_f$ gives $\ddot{\mathbf{x}}_f = \mathbf{A}_f^2 \mathbf{x}_f + \mathbf{A}_f \mathbf{B}_f \mathbf{e}_R + \mathbf{B}_f \dot{\mathbf{e}}_R$, which is bounded. Applying Barbalat's Lemma, $\dot{\mathbf{x}}_f \to \mathbf{0}$. This fact, along with the assumption that $\mathbf{B}_f$ has full rank implies that $\mathbf{e}_R \to \mathbf{0}$ and $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) \to 0$. Consequently, $\mathbf{u}_R \to \mathbf{0}$ as $t \to \infty$. Since the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is the point $\tilde{X}_R = I$, it follows that $\tilde{X}_R \to I$ as $t \to \infty$. Due to the boundedness of $\tilde{X}_R$, $\mathbf{x}_f$, and $\tilde{\mathbf{x}}_d$ as well as the assumption that X and v are bounded it can be shown that $\dddot{\tilde{X}}_R$ is bounded. Thus $\ddot{\tilde{X}}_R$ is uniformly continuous and by Barbalat's Lemma $\ddot{\tilde{X}}_R \to 0$ as $t \to \infty$. From (6.33),

$$\begin{aligned} 0 &= \lim_{t \to \infty} \left(\tilde{X}_R \mathrm{Ad}_X(S(\tilde{\mathbf{d}})) - \mathbf{S}(\mathbf{u}_R) \tilde{X}_R \right) \\ &= \lim_{t \to \infty} \mathrm{Ad}_X(S(\tilde{\mathbf{d}})) \end{aligned}$$

and thus $\tilde{\mathbf{d}} \to \mathbf{0}$ as $t \to \infty$. With the assumption that $\mathbf{C}_d$ has full rank $\tilde{\mathbf{d}} \to \mathbf{0}$ implies that $\tilde{\mathbf{x}}_d \to \mathbf{0}$ as $t \to \infty$.

It has been shown that for all $(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) \in \mathscr{L}_\alpha$, $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \to (\mathbf{1}, \mathbf{0}, \mathbf{0})$ as $t \to \infty$. As V_3 is positive definite on $\mathscr{L}_\alpha$ there exists a class $\mathcal{K}$ function ϕ such that $\phi(\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)) \leq V(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ for all $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \mathscr{L}_\alpha$ [94]. Thus, $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \to (I, \mathbf{0}, \mathbf{0})$ as $t \to \infty$ for all $\ell(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) < \phi^{-1}(\alpha)$ and therefore the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is uniformly convergent [95]. This, along with the fact that the equilibrium point is uniformly stable, shows that $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is locally uniformly asymptotically stable.

To show item (ii) of Theorem 6.12, recall that $\dot{V}_3 \leq 0$ and consequently

$$V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) \leq V_3(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0))$$

for all $t \geq 0$. By assumption $V_3(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) < L$, and therefore

$$V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) < L$$

for all $t \geq 0$. This implies that $\mathbf{x}_f$ and $\tilde{\mathbf{x}}_d$ remain bounded for all $t \geq 0$ and that (6.26) is satisfied for all $t \geq 0$, which implies that $\mathbf{e}_R$ is bounded for all $t \geq 0$. Applying Barbalat's Lemma on V_3, $\mathbf{x}_f$, and $\tilde{X}_R$, as above, it can be shown that $\mathbf{e}_R \to \mathbf{0}$, $\mathbf{x}_f \to \mathbf{0}$, and $\tilde{\mathbf{x}}_d \to \mathbf{0}$ as $t \to \infty$. Due to the fact that $V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) < L$, $\mathbf{e}_R \to \mathbf{0}$ implies that $\tilde{X}_R \to I$. Therefore trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ asymptotically approach $(I, \mathbf{0}, \mathbf{0})$. □

As was the case for Theorem 6.8, the restriction on initial conditions in item (ii) can be interpreted as an estimate of the region of attraction of the equilibrium point $(\tilde{\mathbf{X}}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$. This estimate is given by the set $\{(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in G \times \mathbb{R}^{n_f} \times \mathbb{R}^{n_d} \mid V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \leq c\}$, where $c < L$. It is also the case that if $L = +\infty$, then the equilibrium point is globally asymptotically stable.

6.5.3.2 Stability of the Left Observer with Disturbance Rejection

Theorem 6.13. *Consider trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ under (6.35). Let f_N be a bi-invariant error function about X and take a bi-invariant Riemannian metric. Let $(\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f)$ be SPR, where $\mathbf{B}_f$ has full rank and $\mathbf{D}_f \geq 0$. Assume the following:*

(A1) there is no noise associated with the group velocity measurement or the partial state measurements such that $d = 0$ and $N = \mathcal{I}$;

(A2) there exists a faithful representation of the Lie group G as a matrix Lie group;

(A3) $\mathbf{\Phi}(X)$ *and* v *are bounded with respect to* $||\cdot||_{\mathsf{F}}$ *and* $||\cdot||_{\mathfrak{g}}$*, respectively;*

(A4) $\mathbf{C}_d$ *has full rank.*

Define $L = \sup\{c \in \mathbb{R} \mid \tilde{X}_R \in \Omega_c \setminus \{I\} \implies \mathbf{e}_L \neq \mathbf{0}\}$*, where* $\Omega_c = \{\tilde{X}_R \in G \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq c\}$*. Then the equilibrium point* $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ *is locally uniformly asymptotically stable.*

Proof. The proof follows in a similar manner to the proof of Theorem 5.1 in [16]. Consider the Lyapunov function candidate

$$V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = f_{\mathcal{I}}(\tilde{X}_R, I) + \mathcal{L}(\mathbf{x}_f) + \tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d.$$

The derivative with respect to time of the third term, $\frac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d$, is

$$\frac{\mathrm{d}}{\mathrm{d}t}\tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d = \tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}(\mathbf{A}_d^{\mathsf{T}} + \mathbf{A}_d)\tilde{\mathbf{x}}_d - \tilde{\mathbf{x}}_d^{\mathsf{T}}\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L.$$

Recall that $\mathbf{A}_d$ is skew-symmetric and therefore $\mathbf{A}_d^{\mathsf{T}} + \mathbf{A}_d = \mathbf{0}$. Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\tfrac{1}{2}\rho^{-1}\tilde{\mathbf{x}}_d^{\mathsf{T}}\tilde{\mathbf{x}}_d = -\tilde{\mathbf{x}}_d^{\mathsf{T}}\mathbf{C}_d^{\mathsf{T}}\mathbf{e}_L = -\tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L.$$

The derivative with respect to time of V_3 can therefore be written as

$$\dot{V}_3 = \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L. \tag{6.38}$$

Substituting (6.35a) into (6.38) yields

$$\begin{aligned}
\dot{V}_3 &= \langle \nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I), \dot{\tilde{X}}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L \\
&= \langle \mathrm{Ad}_{\tilde{X}}(e_L)\tilde{X}_R, \tilde{X}_R \mathrm{Ad}_X(\tilde{d} - u_L) \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L \\
&= \langle \mathrm{Ad}_{\tilde{X}}(e_L)\tilde{X}_R, \mathrm{Ad}_{\tilde{X}}(\tilde{d} - u_L)\tilde{X}_R \rangle_{\tilde{X}_R} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L \\
&= \langle S(\mathbf{e}_L), S(\tilde{\mathbf{d}} - \mathbf{u}_L) \rangle + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L \\
&= -\mathbf{e}_L^{\mathsf{T}}\mathbf{u}_L + \mathbf{e}_L^{\mathsf{T}}\tilde{\mathbf{d}} + \dot{\mathcal{L}}(\mathbf{x}_f) - \tilde{\mathbf{d}}^{\mathsf{T}}\mathbf{e}_L \\
&= -\mathbf{e}_L^{\mathsf{T}}\mathbf{u}_L + \dot{\mathcal{L}}(\mathbf{x}_f).
\end{aligned}$$

By Corollary 6.6, $\dot{V}_3 = -\frac{1}{2}\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\mathbf{x}_f - \mathbf{e}_L^{\mathsf{T}}\mathbf{D}_f\mathbf{e}_L$. Therefore, $\dot{V}_3 \leq 0$. As $f_{\mathcal{I}}$ is an error function it follows that $f_{\mathcal{I}}$ is locally positive definite and $\tilde{X}_R = I$ is an isolated critical point of $f_{\mathcal{I}}$ [17]. By assumption, there exists a faithful representation of G as a matrix Lie group. This implies that there exists $m > 0$ and a mapping $\mathbf{\Phi} : G \to GL(m)$ such

that $\mathbf{\Phi}(G)$ is a matrix Lie group [16]. Following the proof of Theorem 5.1 in [16], this implies that there exists a set $B_r = \{\tilde{X}_R \in G \mid d(\tilde{X}_R) \leq r\}$ about $\tilde{X}_R = I$, where $d(\tilde{X}_R) = ||\mathbf{1} - \mathbf{\Phi}(\tilde{X}_R)||_{\mathsf{F}}$, such that for all $\tilde{X}_R \in B_r$, $f_{\mathcal{I}}(\tilde{X}_R, I)$ is positive definite and $\tilde{X}_R = I$ is the only critical point of $f_{\mathcal{I}}$ in B_r. Further, this implies that $V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}})$ is positive definite in the set $\bar{B}_r = \{(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in G \times \mathbb{R}^{n_f} \times \mathbb{R}^{n_d} \mid \ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \leq r\}$, where $\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = d(\tilde{X}_R) + ||\mathbf{x}_f||_2 + ||\tilde{\mathbf{x}}_d||_2$, and $\tilde{X}_R = I$ is the only critical point of $f_{\mathcal{I}}$ in the set $\bar{B}_r$. Consequently, for all $t \geq 0$ and all $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \bar{B}_r$ the Lyapunov function V_3 is positive definite and $\dot{V}_3 \leq 0$ and by Theorem 4.8 of [94] the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is uniformly stable.

Again, following the proof of Theorem 5.1 in [16], choose

$$\alpha < \min_{\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = r} V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d),$$

and define $\mathscr{L}_\alpha = \{(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \bar{B}_r \mid V_3(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \leq \alpha\}$. Then, $\mathscr{L}_\alpha \subset \bar{B}_r$ and all trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ starting in $\mathscr{L}_\alpha$ remain in $\mathscr{L}_\alpha$ for all $t \geq 0$ [94]. Consequently, trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ remain bounded with respect to $\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$, which implies $\tilde{X}_R$ remains bounded with respect to $d(\cdot)$ and $\mathbf{x}_f, \tilde{\mathbf{x}}_d$ remain bounded with respect to $||\cdot||_2$.

As $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \mathscr{L}_\alpha$ for all $t \geq 0$, it follows that $\tilde{X}_R \in \Omega_\alpha$ for all $t \geq 0$, where $\Omega_\alpha = \{X \in B_r \mid f_{\mathcal{I}}(\tilde{X}_R, I) \leq \alpha\}$. Moreover, $\mathscr{L}_\alpha \subset \bar{B}_r$ implies that the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is $\tilde{X}_R = I$ and therefore the only critical point in Ω_α is $\tilde{X}_R = I$. Consequently, $\Omega_\alpha \subset \Omega_L$ and therefore by Proposition 6.30 of [17] there exists constants b_1 and b_2 such that (6.26) is satisfied. This implies that $\mathbf{e}_R$ is bounded for all $t \geq 0$.

Taking the second derivative with respect to time of V_1 yields $\ddot{V}_1 = -\mathbf{x}_f^{\mathsf{T}}\mathbf{Q}_f\dot{\mathbf{x}}_f - 2\mathbf{e}_L^{\mathsf{T}}\mathbf{D}_f\dot{\mathbf{e}}_L$. By Corollary 6.7, $\dot{\mathbf{e}}_R = \mathbf{H}(\tilde{X}_R)\mathbf{q} - \boldsymbol{\xi}$ where $S(\mathbf{q}) = \mathrm{Ad}_{\hat{X}}(\tilde{d} - u_L)$ and $S(\boldsymbol{\xi}) = Q(\mathrm{Ad}_{\hat{X}}(\tilde{d} - u_L), e_R)$. By assumption $||\mathbf{\Phi}(X)||_{\mathsf{F}}$ is bounded, which implies that $||\mathbf{\Phi}(\hat{X})||_{\mathsf{F}}$ is bounded due to the fact that $\tilde{X}_R$ is bounded. Given that $\tilde{X}_R$, $\hat{X}$, $\mathbf{x}_f$, $\tilde{\mathbf{x}}_d$, and $\mathbf{e}_R$ are bounded it follows that $\mathbf{e}_L$, $\mathbf{u}_f$, $\dot{\mathbf{x}}_f$ and $\dot{\mathbf{e}}_R$ are bounded. Recall, $e_L = \mathrm{Ad}_{\hat{X}^{-1}}(e_R)$ and therefore $\dot{e}_L = \mathrm{Ad}_{\hat{X}^{-1}}(\dot{e}_R) + [\mathrm{Ad}_{\hat{X}^{-1}}(e_R), v + \tilde{d} - u_L]$ [17, p. 307]. As v is bounded by assumption, it follows that $\dot{e}_L$ is bounded and thus $\dot{\mathbf{e}}_L$ is bounded as well. Thus, $\ddot{V}_1$ is bounded and therefore $\dot{V}_1$ is uniformly continuous. By application of Barbalat's Lemma $\dot{V}_1 \to 0$ as $t \to \infty$ and thus $\mathbf{x}_f \to \mathbf{0}$ as $t \to \infty$. Taking the derivative with respect to time of $\dot{\mathbf{x}}_f$ gives $\ddot{\mathbf{x}}_f = \mathbf{A}_f^2\mathbf{x}_f + \mathbf{A}_f\mathbf{B}_f\mathbf{e}_L + \mathbf{B}_f\dot{\mathbf{e}}_L$, which is bounded. Applying Barbalat's Lemma, $\dot{\mathbf{x}}_f \to \mathbf{0}$. This fact, along with the assumption that $\mathbf{B}_f$ has full rank implies that $\mathbf{e}_L \to \mathbf{0}$ and $\nabla_{\tilde{X}_R} f_{\mathcal{I}}(\tilde{X}_R, I) \to 0$. Consequently, $\mathbf{u}_L \to \mathbf{0}$ as $t \to \infty$. Since the only critical point of $f_{\mathcal{I}}$ in $\mathscr{L}_\alpha$ is the

point $\tilde{X}_R = I$, it follows that $\tilde{X}_R \to I$ as $t \to \infty$. Due to the boundedness of $\tilde{X}_R$, $\mathbf{x}_f$, and $\tilde{\mathbf{x}}_d$ as well as the assumption that X and v are bounded it can be shown that $\ddot{\tilde{X}}_R$ is bounded. Thus $\dot{\tilde{X}}_R$ is uniformly continuous and by Barbalat's Lemma $\dot{\tilde{X}}_R \to 0$ as $t \to \infty$. From (6.35),

$$\begin{aligned} 0 &= \lim_{t\to\infty} \left(\tilde{X}_R \mathrm{Ad}_X(S(\tilde{\mathbf{d}} - \mathbf{u}_L)) \right) \\ &= \lim_{t\to\infty} \mathrm{Ad}_X(S(\tilde{\mathbf{d}})) \end{aligned}$$

and thus $\tilde{\mathbf{d}} \to \mathbf{0}$ as $t \to \infty$. With the assumption that $\mathbf{C}_d$ has full rank $\tilde{\mathbf{d}} \to \mathbf{0}$ implies that $\tilde{\mathbf{x}}_d \to \mathbf{0}$ as $t \to \infty$.

It has been shown that for all $(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) \in \mathscr{L}_\alpha$, $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \to (\mathbf{1}, \mathbf{0}, \mathbf{0})$ as $t \to \infty$. As V_3 is positive definite on $\mathscr{L}_\alpha$ there exists a class $\mathcal{K}$ function ϕ such that $\phi(\ell(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)) \leq V(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ for all $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \in \mathscr{L}_\alpha$ [94]. Thus, $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) \to (I, \mathbf{0}, \mathbf{0})$ as $t \to \infty$ for all $\ell(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) < \phi^{-1}(\alpha)$ and therefore the equilibrium point $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is uniformly convergent [95]. This, along with the fact that the equilibrium point is uniformly stable, shows that $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d) = (I, \mathbf{0}, \mathbf{0})$ is locally uniformly asymptotically stable.

To show item (ii) of Theorem 6.12, recall that $\dot{V}_3 \leq 0$ and consequently

$$V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) \leq V_3(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0))$$

for all $t \geq 0$. By assumption $V_3(\tilde{X}_R(0), \mathbf{x}_f(0), \tilde{\mathbf{x}}_d(0)) < L$, and therefore

$$V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) < L$$

for all $t \geq 0$. This implies that $\mathbf{x}_f$ and $\tilde{\mathbf{x}}_d$ remain bounded for all $t \geq 0$ and that (6.26) is satisfied for all $t \geq 0$, which implies that $\mathbf{e}_R$ is bounded for all $t \geq 0$. Applying Barbalat's Lemma on V_3, $\mathbf{x}_f$, and $\tilde{X}_R$, as above, it can be shown that $\mathbf{e}_R \to \mathbf{0}$, $\mathbf{x}_f \to \mathbf{0}$, and $\tilde{\mathbf{x}}_d \to \mathbf{0}$ as $t \to \infty$. Due to the fact that $V_3(\tilde{X}_R(t), \mathbf{x}_f(t), \tilde{\mathbf{x}}_d(t)) < L$, $\mathbf{e}_R \to \mathbf{0}$ implies that $\tilde{X}_R \to I$. Therefore trajectories of $(\tilde{X}_R, \mathbf{x}_f, \tilde{\mathbf{x}}_d)$ asymptotically approach $(I, \mathbf{0}, \mathbf{0})$. □

6.6 Closing Remarks

This chapter discusses nonlinear observer design for systems evolving on Lie groups. The proposed methods are a generalization of the gradient-based Lie group observer first proposed in [14]. The observers have several desirable properties. First,

like many recently developed nonlinear observers, the proposed methods are provably locally asymptotically stable about the point at which the state estimate is equal to the true state. Secondly, the observers evolve directly on the underlying Lie group and thus captures the full nonlinear system dynamics. Third, a disturbance observer may be used to reject constant and harmonic disturbances in the velocity measurement. Finally, the introduction of an LTI system acting on the gradient of a cost function allows for greater design freedom when compared to similar observers in the literature. In fact, the introduction of the LTI system has completed the extension of linear complementary filters to nonlinear observer design on Lie groups.

Stability of the proposed nonlinear observers about the desired equilibrium point was shown, provided the cost function satisfied certain invariance properties. These properties are restrictive in practice and future work will focus on relaxing these restrictions. While the stability proofs required invariance of the cost function, the observers themselves can be implemented regardless of the invariance properties of the cost. In the next chapter the optimal design of the LTI system will be considered.

Chapter 7

$\mathcal{H}_2$-Optimal Observer Design

7.1 Introduction

Optimal control is the process of designing a controller for a dynamic system while meeting some optimality condition. Often studied in the literature is the $\mathcal{H}_2$-optimal control problem, which seeks to find a controller that minimizes the $\mathcal{H}_2$-norm of the closed-loop system. When posed in a stochastic framework, the solution to the $\mathcal{H}_2$-optimal observer design problem results in the popular Kalman filter [98]. In the deterministic setting, the $\mathcal{H}_2$-optimal observer is the Luenberger observer where the static observer gain is related to the solution of an algebraic Ricatti equation. In the design process for the Kalman filter and the Luenberger observer, however, the observer is assumed to have some a priori structure. Specifically, the observers copy the system dynamics and append a static innovation term that is the error between true and predicted observations scaled by a static observer gain. The observer gain is static in the sense that there are no dynamics associated with the innovation. This is opposed to dynamic observers where there exists dynamics in the innovation such that the observer gain is dependent on the frequency of the input. Dynamic observers that evolve in Euclidean space are considered in [33–37].

In Chapter 6, the gradient-based observer was modified to include a linear time-invariant system. Returning to the Luenberger observer analogy discussed in Chapter 3, the gradient-based observer can be interpreted as a static observer on a Lie group. Similarly, the proposed observers in Chapter 6 are analogous to dynamic observers on Lie groups, due to the introduction of the linear operator $\boldsymbol{\mathcal{H}}$. As there exists a duality between the optimal design of dynamic observers and the optimal design of dynamic feedback controllers, optimal control methods can be applied to the dynamic observer

design problem. Consequently, in this chapter the problem of optimal dynamic observer design is considered with the goal of developing a filter synthesis method for the Lie group observers proposed in Chapter 6. The synthesis methods presented in this chapter involve linearizing the nonlinear error dynamics associated with the observers presented in Chapter 6. The optimal filter, $\boldsymbol{\mathcal{H}}$, is then found by minimizing the $\mathcal{H}_2$ norm of the closed-loop system, which will have the effect of mitigating the impact of noise on the estimation error. Moreover, the dynamic nature of the proposed Lie group observer allows for the introduction of exogenous signal dynamics in the feedback loop. This is highly desirable as it allows for disturbance rejection by the internal model principle, which states that disturbance rejection can be achieved by incorporating copies of the exogenous dynamics in the feedback loop [99]. As an alternative to the disturbance observer presented in Section 6.5, in this chapter disturbance rejection is explored by incorporating exogenous dynamics directly in the linear filter $\boldsymbol{\mathcal{H}}$.

The remainder of this chapter proceeds with a review of the standard control problem as well as $\mathcal{H}_2$-optimal control in Sec. 7.2. Optimal filter design is considered in Sec. 7.3, while optima filter design with an internal model of disturbances is considered in Sec. 7.4. Finally, closing remarks are given in Sec. 7.6.

7.2 Optimal Control and the Standard Control Problem

Consider the LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}_1\mathbf{w} + \mathbf{B}_2\mathbf{u}, \tag{7.1a}$$

$$\mathbf{z} = \mathbf{C}_1\mathbf{x} + \mathbf{D}_{11}\mathbf{w} + \mathbf{D}_{12}\mathbf{u}, \tag{7.1b}$$

$$\mathbf{y} = \mathbf{C}_2\mathbf{x} + \mathbf{D}_{21}\mathbf{w} + \mathbf{D}_{22}\mathbf{u}, \tag{7.1c}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ is the system state, $\mathbf{z} \in \mathbb{R}^{n_z}$ is the performance signal, $\mathbf{y} \in \mathbb{R}^{n_y}$ is the system output, $\mathbf{w} \in \mathbb{R}^{n_w}$ is the disturbance or exogenous signal, and $\mathbf{u} \in \mathbb{R}^{n_u}$ is the control input. In the frequency domain, (7.1) is often written as

$$\begin{bmatrix} \mathbf{z}(s) \\ \mathbf{y}(s) \end{bmatrix} = \mathbf{G}(s) \begin{bmatrix} \mathbf{w}(s) \\ \mathbf{u}(s) \end{bmatrix}. \tag{7.2}$$

Now, consider the feedback interconnection of the system (7.1) with a controller $\mathbf{G}_c(s)$, via $\mathbf{u}(s) = \mathbf{G}_c(s)\mathbf{y}(s)$, shown in Fig. 7.1. The closed-loop transfer matrix from $\mathbf{w}$ to $\mathbf{z}$ is denoted $\mathbf{G}_{zw}(s)$ such that $\mathbf{z}(s) = \mathbf{G}_{zw}(s)\mathbf{w}(s)$. According to [40], the standard

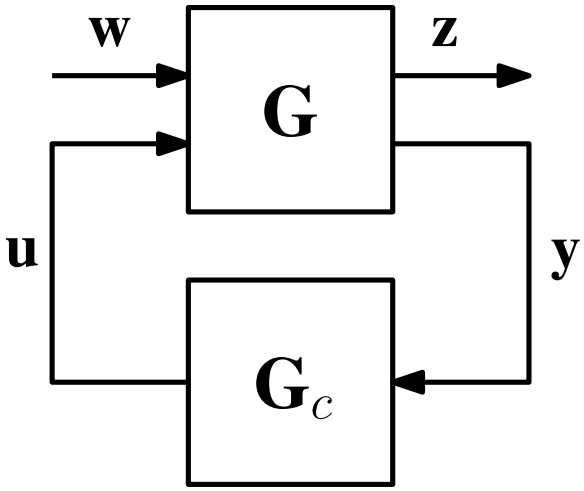


Figure 7.1: The standard control problem.

control problem can be stated as follows: given knowledge of $\mathbf{w}$, determine $\mathbf{G}_c(s)$ such that some performance criterion is minimized. When the performance criterion is the $\mathcal{H}_2$ norm of $\mathbf{G}_{zw}(s)$ the control problem is the $\mathcal{H}_2$-control problem stated formally below.

Problem 7.1 (The $\mathcal{H}_2$-Control Problem [100])**.** The $\mathcal{H}_2$-control problem is to find a proper, real rational controller, $\mathbf{G}_c(s)$, that stabilizes $\mathbf{G}(s)$ internally and minimizes the $\mathcal{H}_2$ norm of the closed-loop transfer function $\mathbf{G}_{zw}(s)$.

The $\mathcal{H}_2$ control problem can be understood as minimizing the effect of exogenous signals on the performance variable. Specifically, minimizing the 2-norm of the performance variable $\mathbf{z}$ given unit impulse exogenous signals. Alternatively, $\mathcal{H}_2$-optimal control can be interpreted from a stochastic point of view. In this interpretation, minimizing the $\mathcal{H}_2$ norm corresponds to minimizing the power of the output signal given a unit intensity white noise input [101].

7.3 Optimal Observer Design

In this section, we will investigate methods to find an optimal filter $\mathbf{H}(s)$ for the nonlinear complementary filter introduced in Chapter 6. To do this, the $\mathcal{H}_2$ control design methodology described above will be applied. This requires the following steps:

1. identification of the nonlinear plant;

2. linearization the nonlinear plant about the equilibrium point;

3. formulation of the a standard control problem;

4. synthesis of the optimal filter, $\mathbf{H}(s)$.

7.3.1 Identification of the Nonlinear Plant

Before any synthesis of an optimal filter $\mathbf{H}(s)$ can take place it is necessary to first identify the nonlinear plant as well as the plant output. Filter design will be considered for both the right and left observers presented in Chapter 6. As such, two nonlinear plants are considered, one associated with the right observer and one associated with the left observer. The nonlinear plants are taken as the error dynamics associated with the observers, which were found in Propositions 6.1 and 6.2. The error dynamics are chosen as this allows for the linearization of the plant about the point $\hat{X} = X$. In the standard control problem the output of the plant is taken as the input to the controller. Consequently, the inputs e_R and e_L are taken as the outputs of the nonlinear plant. For convenience, the error dynamics from (6.14) and (6.16) are reproduced here as

$$\begin{aligned} \dot{\tilde{X}}_R &= \tilde{X}_R \mathrm{Ad}_X(d) - u_R \tilde{X}_R, && \text{(7.3a)} \\ e_R &= \nabla_{\tilde{X}_R} f_N(\tilde{X}_R, I) \tilde{X}_R^{-1}, && \text{(7.3b)} \end{aligned}$$

and

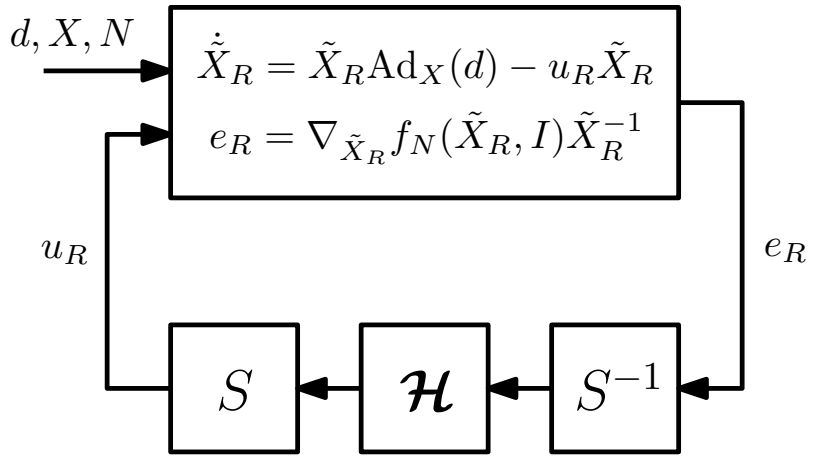
$$\begin{aligned} \dot{\tilde{X}}_L &= -v\tilde{X}_L + \tilde{X}_L v + \tilde{X}_L(d - u_L), && \text{(7.4a)} \\ e_L &= \tilde{X}_L^{-1} \nabla_{\tilde{X}_L} f_N(\tilde{X}_L, I). && \text{(7.4b)} \end{aligned}$$

The optimal filter design problem that is now considered is summarized in Fig. 7.2.

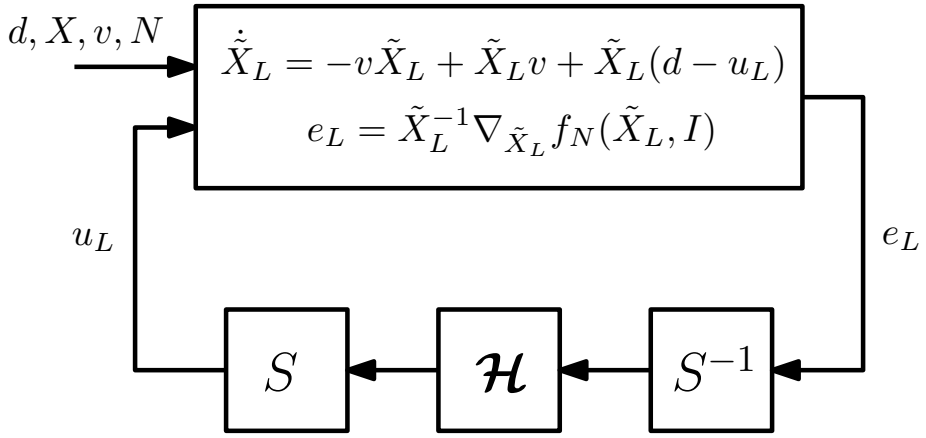
7.3.2 Linearization of the Nonlinear Plant

To design the optimal filter $\mathbf{H}(s)$ for the right and left observers it is necessary to linearize the nonlinear plants in (7.3) and (7.4). The linearization will be taken about the points $(\tilde{X}_R, u_R, d, N) = (I, 0, 0, \mathcal{I})$ and $(\tilde{X}_L, u_L, d, N) = (I, 0, 0, \mathcal{I})$. For this purpose, recall that $d = S(\mathbf{d})$, $u_R = S(\mathbf{u}_R)$, $u_L = S(\mathbf{u}_L)$, and let $\tilde{X}_R = \exp(\tilde{x}_R) = \exp(S(\tilde{\mathbf{x}}_R))$, $\tilde{X}_L = \exp(\tilde{x}_L) = \exp(S(\tilde{\mathbf{x}}_L))$, and

$$N = (\exp(S(\mathbf{n}_1)), \ldots, \exp(S(\mathbf{n}_\ell))), \qquad \text{(7.5)}$$



(a) Block diagram of the right observer error dynamics.



(b) Block diagram of the the left observer error dynamics.

Figure 7.2: Block diagrams of the error dynamics of the right and left observers. The goal is to design $\mathcal{H}$, in some optimal sense, to drive e_R or e_L to zero.

where, $\tilde{x}_R, \tilde{x}_L \in \mathfrak{g}$, and $\mathbf{x}_R, \mathbf{x}_L, \mathbf{n}_1, \ldots, \mathbf{n}_\ell \in \mathbb{R}^n$. The exponential coordinates associated with each N_j, $j \in \{1, \ldots, \ell\}$ can be combined together in $\mathbf{n} \in \mathbb{R}^{n\ell}$ such that $\mathbf{n} = \begin{bmatrix} \mathbf{n}_1^\mathsf{T} & \cdots & \mathbf{n}_\ell^\mathsf{T} \end{bmatrix}^\mathsf{T}$. Then, the points $(\tilde{X}_R, u_R, d, N) = (I, 0, 0, \mathcal{I})$ and $(\tilde{X}_L, u_L, d, N) = (I, 0, 0, \mathcal{I})$ can be expressed as $(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, \mathbf{n}) = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ and $(\tilde{\mathbf{x}}_L, \mathbf{u}_L, \mathbf{d}, \mathbf{n}) = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$, respectively.

The linearization of (7.3) is considered in Proposition 7.2 while the linearization of (7.4) is considered in Proposition 7.3.

Proposition 7.2. *Consider the nonlinear system given in* (7.3)*. Consider perturbations of* $(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, \mathbf{n})$ *about the point* $(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ *such that* $\tilde{\mathbf{x}}_R = \delta\tilde{\mathbf{x}}_R$, $\mathbf{u}_R = \delta\mathbf{u}_R$,

$\mathbf{d} = \delta\mathbf{d}$, *and* $\mathbf{n} = \delta\mathbf{n}$. *Then,* (7.3) *can be linearized as*

$$\begin{aligned} \delta\dot{\tilde{\mathbf{x}}}_R &= \mathbf{B}(t)\delta\mathbf{d} - \delta\mathbf{u}_R, && \text{(7.6a)} \\ \delta\mathbf{e}_R &= \mathbf{M}_1\delta\tilde{\mathbf{x}}_R + \mathbf{M}_2\delta\mathbf{n}. && \text{(7.6b)} \end{aligned}$$

where $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *are appropriately dimensioned constant matrices, and* $\mathbf{B}(t) = [\mathrm{Ad}_{\hat{X}(t)}]_B$.

Proof. The proof follows in a similar manner as the proof of Proposition 3.1 in [102]. First, an expression for $\dot{\tilde{\mathbf{x}}}_R$ will be sought. Recall, by definition $\tilde{X}_R = \exp(S(\tilde{\mathbf{x}}_R))$. Then, by the chain rule

$$\begin{aligned} \dot{\tilde{X}}_R(t) &= T_t\tilde{X}_R(t) \\ &= T_t \exp(S(\tilde{\mathbf{x}}_R(t))) \\ &= T_{S(\tilde{\mathbf{x}}_R(t))} \exp(S(\dot{\tilde{\mathbf{x}}}_R(t))), && \text{(7.7)} \end{aligned}$$

where $T_{S(\tilde{\mathbf{x}}_R)} \exp : \mathfrak{g} \to T_{\exp(S(\tilde{\mathbf{x}}_R))}G$ is the derivative of the exponential map at the point $S(\tilde{\mathbf{x}}_R)$. Substituting (7.7) into (7.3a) it follows that

$$\begin{aligned} T_{S(\tilde{\mathbf{x}}_R)} \exp(S(\dot{\tilde{\mathbf{x}}}_R)) &= \mathrm{Ad}_{\hat{X}}(S(\mathbf{d})) \exp(S(\tilde{\mathbf{x}}_R)) - S(\mathbf{u}_R) \exp(S(\tilde{\mathbf{x}}_R)) \\ &= [\mathrm{Ad}_{\hat{X}}(S(\mathbf{d})) - S(\mathbf{u}_R)] \exp(S(\tilde{\mathbf{x}}_R)). \end{aligned}$$

Inverting the map $T_{S(\tilde{\mathbf{x}}_R)} \exp(\cdot)$ gives

$$S(\dot{\tilde{\mathbf{x}}}_R) = T_{S(\tilde{\mathbf{x}}_R)} \exp^{-1} \left([\mathrm{Ad}_{\hat{X}}(S(\mathbf{d})) - S(\mathbf{u}_R)] \exp(S(\tilde{\mathbf{x}}_R)) \right), \qquad \text{(7.8)}$$

which by Proposition 3.2 in [102] can be expressed as

$$S(\dot{\tilde{\mathbf{x}}}_R) = M_{S(\tilde{\mathbf{x}}_R)}(\mathrm{Ad}_{\hat{X}}(S(\mathbf{d})) - S(\mathbf{u}_R)), \qquad \text{(7.9)}$$

where $M_{S(\tilde{\mathbf{x}}_R)} : \mathfrak{g} \to \mathfrak{g}$ is a mapping such that $M_{S(\tilde{\mathbf{x}}_R)}(a) = T_{S(\tilde{\mathbf{x}}_R)} \exp^{-1}(a \exp(S(\tilde{\mathbf{x}}_R)))$ for all $a \in \mathfrak{g}$ and M_0 is the identity map [102]. Expressing (7.9) in the basis B gives

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}_R &= \mathbf{M}(\tilde{\mathbf{x}}_R)\mathbf{c}(\mathbf{u}_R, \mathbf{d}, t) && \text{(7.10)} \\ &= \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t), && \text{(7.11)} \end{aligned}$$

where $\mathbf{c}(\mathbf{u}_R, \mathbf{d}, t) = \mathbf{B}(t)\mathbf{d} - \mathbf{u}_R$, $\mathbf{M}(\tilde{\mathbf{x}}_R) = [M_{S(\tilde{\mathbf{x}}_R)}]_B$, and $\mathbf{B}(t) = [\mathrm{Ad}_{\hat{X}}]_B$ [102]. Taking

a Taylor series expansion of $\mathbf{f}$ about $(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$ gives [94]

$$\delta\dot{\tilde{\mathbf{x}}}_R \approx \left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \tilde{\mathbf{x}}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} \delta\tilde{\mathbf{x}}_R + \left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \mathbf{u}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} \delta\mathbf{u}_R + \left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \mathbf{d}}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} \delta\mathbf{d}. \tag{7.12}$$

It remains to find expressions for the partial derivatives in (7.12). Starting with the first term,

$$\left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \tilde{\mathbf{x}}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} = \left.\frac{\partial}{\partial \tilde{\mathbf{x}}_R}\mathbf{M}(\tilde{\mathbf{x}}_R)\right|_{\tilde{\mathbf{x}}_R=\mathbf{0}} \mathbf{c}(\mathbf{0}, \mathbf{0}, t), \tag{7.13}$$

where the multiplication operation in (7.13) is done element wise. Due to the fact that $\mathbf{c}(\mathbf{0}, \mathbf{0}, t) = \mathbf{0}$,

$$\left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \tilde{\mathbf{x}}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} = \mathbf{0}. \tag{7.14}$$

Next,

$$\begin{aligned} \left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \mathbf{u}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} &= \mathbf{M}(\mathbf{0}) \left.\frac{\partial}{\partial \mathbf{u}_R}\mathbf{c}(\mathbf{u}_R, \mathbf{d}, t)\right|_{\mathbf{u}_R=\mathbf{d}=\mathbf{0}} && (7.15)\\ &= -\mathbf{M}(\mathbf{0}). && (7.16) \end{aligned}$$

Recall, M_0 is the identity mapping and therefore $\mathbf{M}(\mathbf{0}) = \mathbf{1}$ and consequently,

$$\left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \mathbf{u}_R}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} = -\mathbf{1}. \tag{7.17}$$

Finally,

$$\begin{aligned} \left.\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}_R, \mathbf{u}_R, \mathbf{d}, t)}{\partial \mathbf{d}}\right|_{(\mathbf{0},\mathbf{0},\mathbf{0})} &= \mathbf{M}(\mathbf{0}) \left.\frac{\partial}{\partial \mathbf{d}}\mathbf{c}(\mathbf{u}_R, \mathbf{d}, t)\right|_{\mathbf{u}_R=\mathbf{d}=\mathbf{0}} && (7.18)\\ &= \mathbf{M}(\mathbf{0})\mathbf{B}(t) && (7.19)\\ &= \mathbf{B}(t). && (7.20) \end{aligned}$$

Combining the above shows that $\delta\dot{\tilde{\mathbf{x}}}_R = \mathbf{B}(t)\delta\mathbf{d} - \delta\mathbf{u}_R$. To show (7.6b), recall that

$$e_R = \nabla_{\tilde{X}_R} f(\tilde{X}_R, I, N)\tilde{X}_R^{-1}. \tag{7.21}$$

As e_R depends only on $\tilde{X}_R$ and N, this can be rewritten as $e_R = H(\tilde{X}_R, N)$, where $H : G \times G^\ell \to \mathfrak{g}$. Expressing $\tilde{X}_R$ as $\tilde{X}_R = \exp(S(\tilde{\mathbf{x}}_R))$ and N as in (7.5) gives $e_R = H(\exp(S(\tilde{\mathbf{x}}_R)), (\exp(S(\tilde{\mathbf{n}}_1)), \ldots, \exp(S(\mathbf{n}_\ell))))$, which may be redefined as $e_R =$

$h(S(\tilde{\mathbf{x}}_R), S(\mathbf{n}_1), \ldots, S(\mathbf{n}_\ell))$, where $h : \mathfrak{g} \times \mathfrak{g} \times \cdots \times \mathfrak{g} \to \mathfrak{g}$. It follows then that

$$\mathbf{e}_R = \mathbf{h}(\tilde{\mathbf{x}}_R, \mathbf{n}), \tag{7.22}$$

where $\mathbf{h} : \mathbb{R}^n \times \mathbb{R}^{n\ell}$ is defined as $\mathbf{h}(\tilde{\mathbf{x}}_R, \mathbf{n}) \triangleq S^{-1}(h(S(\tilde{\mathbf{x}}_R), S(\mathbf{n}_1), \ldots, S(\mathbf{n}_\ell)))$. A Taylor series expansion of $\mathbf{h}$ about $\tilde{\mathbf{x}}_R = \mathbf{0}$ and $\mathbf{n} = \mathbf{0}$ gives

$$\begin{aligned} \delta\mathbf{e}_R &= \left.\frac{\partial \mathbf{h}(\tilde{\mathbf{x}}_R, \mathbf{n})}{\partial \tilde{\mathbf{x}}_R}\right|_{\tilde{\mathbf{x}}_R=\mathbf{0},\mathbf{n}=\mathbf{0}} \delta\tilde{\mathbf{x}}_R + \left.\frac{\partial \mathbf{h}(\tilde{\mathbf{x}}_R, \mathbf{n})}{\partial \mathbf{n}_R}\right|_{\tilde{\mathbf{x}}_R=\mathbf{0},\mathbf{n}=\mathbf{0}} \delta\mathbf{n} && (7.23) \\ &= \mathbf{M}_1\delta\tilde{\mathbf{x}}_R + \mathbf{M}_2\delta\mathbf{n}, && (7.24) \end{aligned}$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are constant as $\mathbf{h}$ depends only on $\tilde{\mathbf{x}}_R$ and $\mathbf{n}$. $\square$

Proposition 7.3. *Consider the nonlinear system given in* (7.4)*. Consider perturbations of* $(\tilde{\mathbf{x}}_L, \mathbf{u}_L, \mathbf{d}, \mathbf{n})$ *about the point* $(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ *such that* $\tilde{\mathbf{x}}_L = \delta\tilde{\mathbf{x}}_L$*,* $\mathbf{u}_L = \delta\mathbf{u}_L$*,* $\mathbf{d} = \delta\mathbf{d}$*, and* $\mathbf{n} = \delta\mathbf{n}$*. Then,* (7.4) *can be linearized as*

$$\begin{aligned} \delta\dot{\tilde{\mathbf{x}}}_L &= \mathbf{A}(t)\delta\tilde{\mathbf{x}}_L + \delta\mathbf{d} - \delta\mathbf{u}_L, && (7.25a) \\ \delta\mathbf{e}_L &= \mathbf{M}_1\delta\tilde{\mathbf{x}}_L + \mathbf{M}_2\delta\mathbf{n}. && (7.25b) \end{aligned}$$

where $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *are appropriately dimensioned constant matrices, and* $\mathbf{A}(t) = [-\mathrm{ad}_{v(t)}]_B$*.*

Proof. The proof follows in a similar manner as the proof of Proposition 7.2 and the proof of Proposition 3.2 in [102]. As in Proposition 7.2, first an expression for $\dot{\tilde{x}}_L$ is found. By definition $\tilde{X}_L = \exp(S(\tilde{\mathbf{x}}_L))$ and by the chain rule

$$\begin{aligned} \dot{\tilde{X}}_L &= T_t\tilde{X}_L(t) \\ &= T_t \exp(S(\tilde{\mathbf{x}}_L)) && (7.26) \\ &= T_{S(\tilde{\mathbf{x}}_L)} \exp(S(\dot{\tilde{\mathbf{x}}}_L)), && (7.27) \end{aligned}$$

where, again, $T_{S(\tilde{\mathbf{x}}_L)} \exp : \mathfrak{g} \to T_{\exp(S(\tilde{\mathbf{x}}_L))}G$. Substituting (7.27) into (7.4a) it follows that

$$T_{S(\tilde{\mathbf{x}}_L)} \exp(S(\dot{\tilde{\mathbf{x}}}_L)) = \exp(S(\tilde{\mathbf{x}}_L))(-\mathrm{Ad}_{\exp(-S(\tilde{\mathbf{x}}_L))}(v) + v + d - u_L).$$

Inverting the map $T_{S(\tilde{\mathbf{x}}_L)} \exp(\cdot)$ gives

$$S(\dot{\tilde{\mathbf{x}}}_L) = T_{S(\tilde{\mathbf{x}}_L} \exp^{-1}\left(\exp(S(\tilde{\mathbf{x}}_L))\left[-\mathrm{Ad}_{\exp(-S(\tilde{\mathbf{x}}_L))}(v) + v + d - u_L\right]\right), \tag{7.28}$$

which by Proposition 3.2 in [102] can be expressed as

$$S(\dot{\tilde{\mathbf{x}}}_L) \;=\; M_{S(\tilde{\mathbf{x}}_L)}(-\mathrm{Ad}_{\exp(-S(\tilde{\mathbf{x}}_L))}(v) + v + d - u_L), \tag{7.29}$$

where $M_{S(\tilde{\mathbf{x}}_R)} : \mathfrak{g} \to \mathfrak{g}$ is a mapping such that $M_{S(\tilde{\mathbf{x}}_L)}(a) = T_{S(\tilde{\mathbf{x}}_R)} \exp^{-1}(\exp(S(\tilde{\mathbf{x}}_R))a)$ for all $a \in \mathfrak{g}$ and M_0 is the identity map [102]. Define $c : \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g} \times \mathbb{R}$ such that $c(\tilde{x}_L, u_L, d, t) = -\mathrm{Ad}_{\exp(-S(\tilde{\mathbf{x}}_L))}(v) + v + d - u_L$. Then, (7.29) is expressed as [102]

$$\dot{\tilde{x}}_L = M_{\tilde{x}_L}(c(\tilde{x}_L, u_L, d, t)). \tag{7.30}$$

Let $\mathrm{D}_1 c(0,0,0,t) : \mathfrak{g} \to \mathfrak{g}$, $\mathrm{D}_2 c(0,0,0,t) : \mathfrak{g} \to \mathfrak{g}$, and $\mathrm{D}_3 c(0,0,0,t) : \mathfrak{g} \to \mathfrak{g}$ be the derivatives of the map c at the point $(\tilde{x}_L, u_L, d, t) = (0,0,0,t)$ with respect to the first, second, and third arguments of c respectively. As M_0 is the identity mapping it follows, by Proposition 3.2 in [102], that

$$\delta\dot{\tilde{x}}_L = \mathrm{D}_1 c(0,0,0,t)(\delta\tilde{x}_L) + \mathrm{D}_2 c(0,0,0,t)(\delta u_L) + \mathrm{D}_3 c(0,0,0,t)(\delta d). \tag{7.31}$$

To compute $\mathrm{D}_1 c(0,0,0,t)(\delta\tilde{x}_L)$, define a curve $\gamma : \mathbb{R} \to \mathfrak{g}$, $\gamma(\epsilon) = \epsilon\delta\tilde{x}_L$, such that $\gamma(0) = 0$ and $\gamma'(0) = \delta\tilde{x}_L$. Then, $\mathrm{D}_1 c(0,0,0,t)(\delta\tilde{x}_L)$ satisfies

$$\begin{aligned}
\mathrm{D}_1 c(0,0,0,t)(\delta\tilde{x}_L) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon} c(\gamma(\epsilon), u_L, d, t)\Big|_{\tilde{x}_L = u_L = d = 0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon}(-\mathrm{Ad}_{\exp(-\epsilon\delta\tilde{x}_L)}(v) + v + d - u_L)\Big|_{\tilde{x}_L = u_L = d = 0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon}(-\mathrm{Ad}_{\exp(-\epsilon\delta\tilde{x}_L)}(v))\Big|_{\tilde{x}_L = u_L = d = 0} \\
&= -\mathrm{ad}_{-\delta\tilde{x}_L}(v),
\end{aligned} \tag{7.32}$$

where the relationship $\mathrm{d}/\mathrm{d}\epsilon(\mathrm{Ad}_{\exp(\epsilon x)})|_{\epsilon=0} = \mathrm{ad}_x$, for all $x \in \mathfrak{g}$, has been used [102]. Recall from the definition of ad that $\mathrm{ad}_x(y) = [x, y]$ for all $x, y \in \mathfrak{g}$ and $[\cdot, \cdot]$ is the Lie bracket associated with $\mathfrak{g}$. From the properties of the Lie bracket, it follows that

$$\begin{aligned}
\mathrm{ad}_{-\delta\tilde{x}_L}(v) &= [-\delta\tilde{x}_L, v] \\
&= -[\delta\tilde{x}_L, v] \\
&= [v, \delta\tilde{x}_L] \\
&= \mathrm{ad}_v(\delta\tilde{x}_L).
\end{aligned} \tag{7.33}$$

Following a similar procedure for $\mathrm{D}_2 c(0,0,0,t)(\delta u_L)$ and $\mathrm{D}_3 c(0,0,0,t)(\delta d)$ gives

$$\mathrm{D}_2 c(0,0,0,t)(\delta u_L) = -\delta u_L, \tag{7.34}$$

and

$$\mathrm{D}_3 c(0,0,0,t)(\delta u_L) = \delta d. \tag{7.35}$$

Substituting (7.32), (7.33), (7.34), and (7.35) into (7.31) gives

$$\delta\dot{\tilde{x}}_L = -\mathrm{ad}_v(\delta\tilde{x}_L) - \delta u_L + \delta d. \tag{7.36}$$

Resolving (7.36) in the basis B yields (7.25a). Equation (7.25b) follows in the exact same manner as in the proof of Proposition 7.2. $\square$

The linearization of the error dynamics for both the right and left observers results in linear time-varying (LTV) systems due to the presence of $\mathbf{B}(t)$ in (7.6) and $\mathbf{A}(t)$ in (7.25). The filter synthesis method considered in this chapter requires that the linearized system be time-invariant rather than time-varying. Consequently, for design purposes the LTV systems in (7.6) and (7.25) will be modified in the following way. For the linearization of the right observer given in (7.6), the exogenous signal $\delta\mathbf{d}(t)$ is combined with $\mathbf{B}(t)$ to form a single exogenous signal $\delta\mathbf{d}'(t) = \mathbf{B}(t)\mathbf{d}(t)$. The resulting LTI system is given by

$$\begin{aligned} \delta\dot{\tilde{\mathbf{x}}}_R &= \delta\mathbf{d}' - \delta\mathbf{u}_R, & (7.37a)\\ \delta\mathbf{e}_R &= \mathbf{M}_1\delta\tilde{\mathbf{x}}_R + \mathbf{M}_2\delta\mathbf{n}. & (7.37b) \end{aligned}$$

For the linearization of the left observer given in (7.25), the assumption that $\mathbf{A}(t)\delta\tilde{\mathbf{x}}_L = \mathbf{0}$ is made. Alternatively, $\mathbf{A}(t)\delta\tilde{\mathbf{x}}_L$ may be considered as an exogenous signal and be lumped in with $\delta\mathbf{d}(t)$. In any case, the left observer linearization is simplified to

$$\begin{aligned} \delta\dot{\tilde{\mathbf{x}}}_L &= \delta\mathbf{d} - \delta\mathbf{u}_L, & (7.38a)\\ \delta\mathbf{e}_L &= \mathbf{M}_1\delta\tilde{\mathbf{x}}_L + \mathbf{M}_2\delta\mathbf{n}. & (7.38b) \end{aligned}$$

As (7.37) and (7.38) are of the same form, the remainder of this section will consider systems of the form

$$\begin{aligned} \delta\dot{\tilde{\mathbf{x}}} &= \delta\mathbf{d} - \delta\mathbf{u}, & (7.39a)\\ \delta\mathbf{e} &= \mathbf{M}_1\delta\tilde{\mathbf{x}} + \mathbf{M}_2\delta\mathbf{n}, & (7.39b) \end{aligned}$$

where the subscripts have been dropped.

7.3.3 Formulation of the Standard Control Problem

To employ $\mathcal{H}_2$-optimal control methods, the system in (7.39) will need to be written in a similar form as (7.1). For this purpose, define a nominal plant $\mathbf{G}_p(s)$ given by

$$\delta\dot{\tilde{\mathbf{x}}} = \mathbf{A}_p\delta\tilde{\mathbf{x}} + \mathbf{B}_p(\delta\mathbf{d} - \delta\mathbf{u}), \tag{7.40}$$

$$\mathbf{y}_p = \mathbf{C}_p\delta\tilde{\mathbf{x}}, \tag{7.41}$$

where $\mathbf{y}_p(s) = \mathbf{G}_p(s)(\delta\mathbf{d}'(s) - \delta\mathbf{u}(s))$, $\mathbf{A}_p = \mathbf{0}$, $\mathbf{B}_p = \mathbf{1}$ and $\mathbf{C}_p = \mathbf{1}$. The block diagram of the feedback interconnection between (7.39) and the filter $\mathbf{H}(s)$ is shown in Fig. 7.3. It is now possible to pose the standard control problem to find the optimal

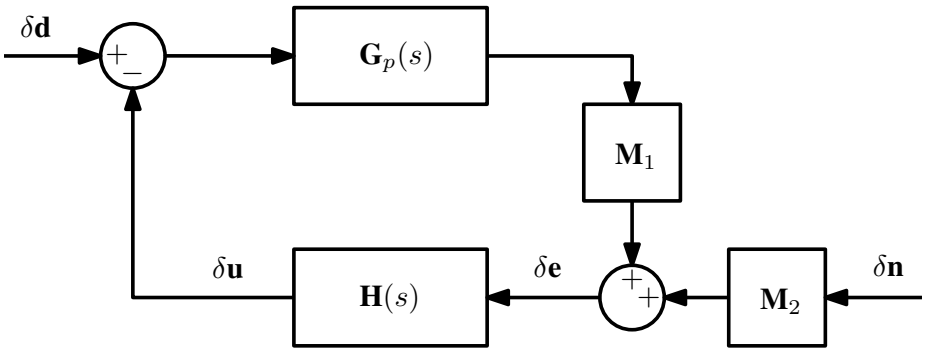


Figure 7.3: Block diagram of interconnection between the nominal plant and the filter $\mathbf{H}(s)$.

filter $\mathbf{H}(s)$ for the plant $\mathbf{G}_p(s)$. However, the plant for which the filter $\mathbf{H}(s)$ is being designed, $\mathbf{G}_p(s)$, contains no useful information relating to the exogenous signals $\delta\mathbf{d}$ and $\delta\mathbf{n}$. Further, as $\mathbf{G}_p(s)$ is simply an integrator the optimal design of $\mathbf{H}(s)$ in this case is not a particularly interesting problem. To make the optimal design of $\mathbf{H}(s)$ for $\mathbf{G}_p(s)$ a useful exercise it is necessary to add weights to the block diagram structure shown in Fig. 7.3. The modified block diagram with added weights is shown in Fig. 7.4 [103]. The weights take the form of LTI systems given by $\mathbf{W}_d(s)$, $\mathbf{W}_n(s)$, $\mathbf{W}_{z_1}(s)$,

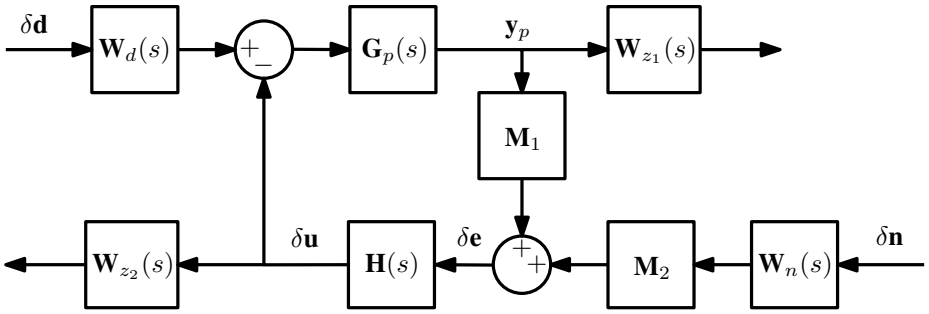


Figure 7.4: Block diagram of interconnection between the nominal plant and the filter $\mathbf{H}(s)$. Frequency dependent weights are added to the exogenous signals and the performance variable. The frequency weights allow the targeting of specific frequency bandwidths during the filter design procedure.

and $\mathbf{W}_{z_2}(s)$ with minimal state-space representations

$$\begin{aligned}
\mathbf{W}_d(s) &\sim \left[\begin{array}{c|c} \mathbf{A}_d & \mathbf{B}_d \\ \hline \mathbf{C}_d & \mathbf{D}_d \end{array}\right], \; \mathbf{W}_n(s) \sim \left[\begin{array}{c|c} \mathbf{A}_n & \mathbf{B}_n \\ \hline \mathbf{C}_n & \mathbf{D}_n \end{array}\right], \\
\mathbf{W}_{z_1}(s) &\sim \left[\begin{array}{c|c} \mathbf{A}_{z_1} & \mathbf{B}_{z_1} \\ \hline \mathbf{C}_{z_1} & \mathbf{D}_{z_1} \end{array}\right], \quad \text{and } \mathbf{W}_{z_2}(s) \sim \left[\begin{array}{c|c} \mathbf{A}_{z_2} & \mathbf{B}_{z_2} \\ \hline \mathbf{C}_{z_2} & \mathbf{D}_{z_2} \end{array}\right].
\end{aligned}$$

This allows for the weighting of certain frequency bandwidths in $\delta\mathbf{d}(s)$ and $\delta\mathbf{n}(s)$, which, after the design process, gives the filter synthesis method information on the frequency content of the exogenous signals. In effect, the weights model the frequency content of the exogenous signals, which is then exploited in the design process. For example, if $\mathbf{d}(s)$ is comprised primarily of low frequency content, then $\mathbf{W}_d(s)$ can be chosen as a low-pass filter. This will have the effect of weighting the low frequency content in $\mathbf{d}(s)$ resulting in better noise rejection at those frequencies. Let $\mathbf{x}_d \in \mathbb{R}^{n_d}$, $\mathbf{x}_n \in \mathbb{R}^{n_d}$, and $\mathbf{x}_z \in \mathbb{R}^{n_z}$ be the states associated with $\mathbf{W}_d(s)$, $\mathbf{W}_n(s)$, and $\mathbf{W}_z(s)$, respectively. Further, define the augmented state variable as $\mathbf{x} = \left[\begin{array}{cccc} \delta\tilde{\mathbf{x}}^\mathsf{T} & \mathbf{x}_d^\mathsf{T} & \mathbf{x}_n^\mathsf{T} & \mathbf{x}_z^\mathsf{T} \end{array}\right]^\mathsf{T}$, the performance variable $\mathbf{z}(s) = \mathbf{W}_{z_1}(s)\mathbf{y}_p(s) + \mathbf{W}_{z_2}(s)\mathbf{u}(s)$, and the exogenous signal as $\mathbf{w} = [\ \delta\mathbf{d}^\mathsf{T} \ \ \delta\mathbf{n}^\mathsf{T}\ ]^\mathsf{T}$, the filter input $\mathbf{y} = \mathbf{C}_p\delta\tilde{\mathbf{x}} + \mathbf{M}_2\delta\mathbf{n}$, and filter output $\mathbf{u} = \delta\mathbf{u}$. Then, we may write

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}_1\mathbf{w} + \mathbf{B}_2\mathbf{u}, \tag{7.42a}$$

$$\mathbf{z} = \mathbf{C}_1\mathbf{x} + \mathbf{D}_{11}\mathbf{w} + \mathbf{D}_{12}\mathbf{u}, \tag{7.42b}$$

$$\mathbf{y} = \mathbf{C}_2\mathbf{x} + \mathbf{D}_{21}\mathbf{w} + \mathbf{D}_{22}\mathbf{u}, \tag{7.42c}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_p & \mathbf{B}_p\mathbf{C}_d & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_d & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_z\mathbf{C}_p & \mathbf{0} & \mathbf{0} & \mathbf{A}_{z_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{z_2} \end{bmatrix}, \ \mathbf{B}_1 = \begin{bmatrix} \mathbf{B}_p\mathbf{D}_d & \mathbf{0} \\ \mathbf{B}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_n \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{B}_2 = \begin{bmatrix} -\mathbf{B}_p \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix},$$

$$\mathbf{C}_1 = \begin{bmatrix} \mathbf{D}_z\mathbf{C}_p & \mathbf{0} & \mathbf{0} & \mathbf{C}_{z_1} & \mathbf{C}_{z_2} \end{bmatrix}, \ \mathbf{D}_{11} = \mathbf{0}, \ \mathbf{D}_{12} = \mathbf{D}_{z_2},$$

and

$$\mathbf{C}_2 = \begin{bmatrix} \mathbf{M}_1 & \mathbf{0} & \mathbf{M}_2 & \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{D}_{21} = \begin{bmatrix} \mathbf{0} & \mathbf{M}_2\mathbf{D}_n \end{bmatrix}, \ \mathbf{D}_{22} = \mathbf{0}$$

7.3.4 Filter Synthesis

The optimal design of $\mathbf{H}(s)$ has been posed as a standard control problem. It is well known that the $\mathcal{H}_2$-optimal filter, denoted $\mathbf{H}_{\mathcal{H}_2}(s)$, has the realization

$$\mathbf{H}_{\mathcal{H}_2}(s) \sim \left[\begin{array}{c|c} \mathbf{A}_{\mathcal{H}_2} & \mathbf{B}_{\mathcal{H}_2} \\ \hline \mathbf{C}_{\mathcal{H}_2} & \mathbf{0} \end{array}\right], \tag{7.43}$$

where $\mathbf{A}_{\mathcal{H}_2} = \mathbf{A} - \mathbf{B}_2\mathbf{K} - \mathbf{L}\mathbf{C}_2$, $\mathbf{B}_{\mathcal{H}_2} = \mathbf{L}$, $\mathbf{C}_{\mathcal{H}_2} = \mathbf{K}$, $\mathbf{K} = (\mathbf{D}_{12}^{\mathsf{T}}\mathbf{D}_{12})^{-1}\mathbf{B}_2^{\mathsf{T}}\mathbf{P}$, and $\mathbf{L} = \mathbf{\Pi}\mathbf{C}_2^{\mathsf{T}}(\mathbf{D}_{21}\mathbf{D}_{21}^{\mathsf{T}})^{-1}$ [100]. The matrices $\mathbf{P} = \mathbf{P}^{\mathsf{T}} > 0$ and $\mathbf{\Pi} = \mathbf{\Pi}^{\mathsf{T}} > 0$ are the solutions to the algebraic Riccati equations

$$\mathbf{P}\mathbf{A} + \mathbf{A}^{\mathsf{T}}\mathbf{P} - \mathbf{P}\mathbf{B}_2(\mathbf{D}_{12}^{\mathsf{T}}\mathbf{D}_{12})^{-1}\mathbf{B}_2^{\mathsf{T}}\mathbf{P} + \mathbf{C}_1^{\mathsf{T}}\mathbf{C}_1 = \mathbf{0}, \tag{7.44}$$

$$\mathbf{\Pi}\mathbf{A}^{\mathsf{T}} + \mathbf{A}\mathbf{\Pi} - \mathbf{\Pi}\mathbf{C}_2^{\mathsf{T}}(\mathbf{D}_{21}\mathbf{D}_{21}^{\mathsf{T}})^{-1}\mathbf{C}_2\mathbf{\Pi} + \mathbf{B}_1\mathbf{B}_1^{\mathsf{T}} = \mathbf{0}. \tag{7.45}$$

However, recall from Theorems 6.8 and 6.9 that asymptotic convergence of the nonlinear complementary filter is guaranteed provided $\mathbf{H}(s)$ is SPR. Consequently, it is desirable to ensure that the resulting optimal filter is SPR. This can be accomplished in a number of ways. Specifically, the synthesis of optimal SPR controllers is considered in [104–112]. In this dissertation, the SPR controller synthesis method discussed in [112] is employed, although any method that guarantees that the resulting controller is SPR can be used. In this method the "controller" that is sought, $\mathbf{H}(s)$, has

the realization

$$\mathbf{H}(s) \sim \left[\begin{array}{c|c} \mathbf{A}_f & \mathbf{B}_f \\ \hline \mathbf{C}_f & \mathbf{0} \end{array}\right], \tag{7.46}$$

where $\mathbf{A}_f = \mathbf{A}_{\mathcal{H}_2}$ and $\mathbf{B}_f = \mathbf{B}_{\mathcal{H}_2}$ are the dynamic matrix and observer gain associated with the nominal $\mathcal{H}_2$-optimal filter given in (7.43). It is assumed that $\mathbf{A}_{\mathcal{H}_2}$ is asymptotically stable. The matrix $\mathbf{C}_f$ is sought such that:

(a) $\mathbf{H}(s)$ is as close as possible to $\mathbf{H}_{\mathcal{H}_2}(s)$;

(b) resulting filter $\mathbf{H}(s)$ is SPR.

In the synthesis procedure, the first objective is accomplished by minimizing the objective function

$$\mathcal{J} = \mathrm{tr}((\mathbf{C}_f - \mathbf{C}_{\mathcal{H}_2})(\mathbf{C}_f - \mathbf{C}_{\mathcal{H}_2})^\mathsf{T}), \tag{7.47}$$

which minimizes the difference between the $\mathcal{H}_2$-optimal feedback gain and $\mathbf{C}_f$. Recall from Lemma 6.5 that $\mathbf{H}(s)$ is SPR if there exists $\mathbf{P}_f = \mathbf{P}_f^\mathsf{T} > 0$ such that

$$\begin{aligned} \mathbf{P}_f\mathbf{A}_f + \mathbf{A}_f^\mathsf{T}\mathbf{P}_f &< 0 && (7.48) \\ \mathbf{P}_f\mathbf{B}_f &= \mathbf{C}_f^\mathsf{T}. && (7.49) \end{aligned}$$

Then, the second objective can be met by enforcing (7.48) and (7.49) during the synthesis procedure. From (7.49), the objective function in (7.47) can be reformulated in terms of $\mathbf{P}_f$ as

$$\mathcal{J} = \mathrm{tr}((\mathbf{B}_f^\mathsf{T}\mathbf{P}_f - \mathbf{C}_{\mathcal{H}_2})(\mathbf{B}_f^\mathsf{T}\mathbf{P}_f - \mathbf{C}_{\mathcal{H}_2})^\mathsf{T}). \tag{7.50}$$

Introducing slack variables $g \in \mathbb{R}^+$ and $\mathbf{Z} \in \mathbb{R}^{n\times n}$, the SPR synthesis method in [112] can be stated as

$$\begin{aligned} \min_{\mathbf{P}_f,\mathbf{Z},g} \quad & \mathcal{J} = g && (7.51a) \\ \text{subject to} \quad & \mathbf{P}_f = \mathbf{P}_f^\mathsf{T} > 0, && (7.51b) \\ & \mathbf{P}_f\mathbf{A}_f + \mathbf{A}_f^\mathsf{T}\mathbf{P}_f < 0, && (7.51c) \\ & \mathbf{Z} = \mathbf{Z}^\mathsf{T} \geq 0, && (7.51d) \\ & \mathrm{tr}(\mathbf{Z}) \leq g, && (7.51e) \\ & \left[\begin{array}{cc} \mathbf{Z} & \mathbf{B}_f^\mathsf{T}\mathbf{P}_f - \mathbf{C}_{\mathcal{H}_2} \\ (\mathbf{B}_f^\mathsf{T}\mathbf{P}_f - \mathbf{C}_{\mathcal{H}_2})^\mathsf{T} & \mathbf{1} \end{array}\right] \geq 0. && (7.51f) \end{aligned}$$

The matrix $\mathbf{C}_f$ is then found after the minimization by (7.49).

7.3.5 Optimal Filter Synthesis Summary

To summarize this section, a review of the synthesis procedure starting from the right or left observer error dynamics in (6.15) or (6.18) is provided.

1. Linearize the nonlinear system (6.15) or (6.18) as in Propositions 7.2 and 7.3.
2. Pick weights $\mathbf{W}_d(s)$, $\mathbf{W}_n(s)$, $\mathbf{W}_{z_1}(s)$, and $\mathbf{W}_{z_2}(s)$.
3. Formulate the standard control problem in (7.42).
4. Find the optimal $\mathcal{H}_2$ feedback controller that stabilizes (7.42).
5. Solve the optimization problem in (7.51) to construct the SPR system $\boldsymbol{\mathcal{H}}$.

7.4 Optimal Observer Design with an Internal Model

In Sec. 6.5 it was shown that the nonlinear complementary filter coupled with a disturbance observer rendered the desired equilibrium point asymptotically stable, provided the linear filter was SPR. The observer equations in this case are given by (6.32) with error dynamics in (6.33). The block diagram for the error dynamics of the right observer with disturbance rejection is shown in Fig. 7.5. In the figure the operator $\boldsymbol{\mathcal{H}}_d$ is the disturbance observer in (6.32d)-(6.32e).

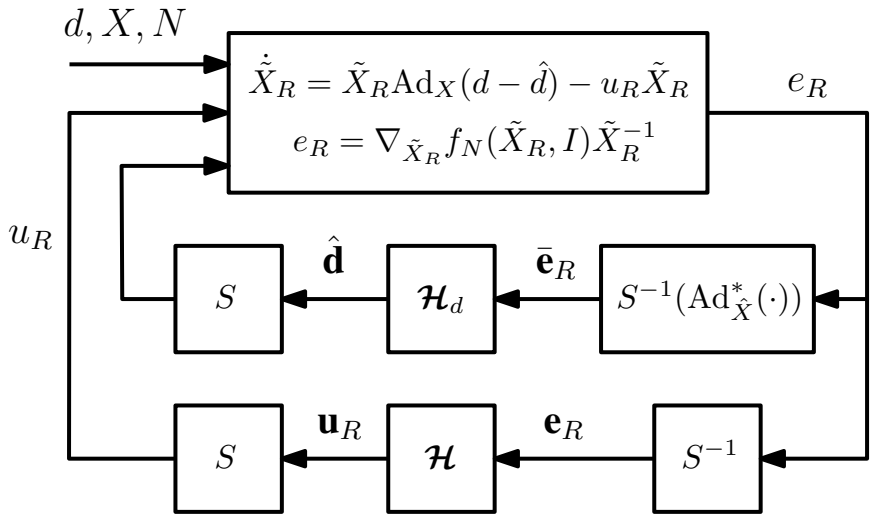


Figure 7.5: Block diagram of the right observer error dynamics with a disturbance observer. The operator $\boldsymbol{\mathcal{H}}$ has no knowledge of the disturbance observer $\boldsymbol{\mathcal{H}}_d$.

In this section, the disturbance rejection problem will be reconsidered in light of the optimal filter design procedure discussed in the previous section. The approach

taken will be that of the internal model controller described in [99]. This method involves augmenting the original LTI system (7.1) with a model of the disturbance. A feedback controller is then designed to stabilize the augmented plant and the final controller implemented is the feedback controller augmented with the internal model. It is shown in [99] that, provided some assumptions are made on the plant, the designed controller is able to reject disturbances and regulate the performance variable $\mathbf{z}$ to zero.

In the previous section, an $\mathcal{H}_2$-optimal filter was designed for both the right and left observers by first linearizing the nonlinear error dynamics and formulating the observer design problem as a standard control problem. Recall from Propositions 7.2 that the linearization of the right observer error dynamics resulted in linear time varying system due to the presence of a time varying input matrix, $\mathbf{B}(t)$, acting on the disturbance $\delta\mathbf{d}$. Previously, the time varying nature of the linearized system was ignored by defining a new exogenous signal $\delta\mathbf{d}'$ and defining the exogenous signals for the standard control problem as $\mathbf{w} = \begin{bmatrix} \delta\mathbf{d}'^{\mathsf{T}} & \delta\mathbf{n}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$. However, in this section it will not be possible to make this simplification as the design method in [99] requires that the input disturbance $\delta\mathbf{d}$ appear directly as an exogenous signal in $\mathbf{w}$. To accommodate this requirement, the observer design with an internal model is restricted to the left observer, whose linearization found in Proposition 7.3 results in a constant input matrix associated with $\delta\mathbf{d}$.

7.4.1 Filter Synthesis

For linear systems of the form (7.1), the disturbance rejection problem can be stated as follows: construct a system $\boldsymbol{\mathcal{H}}$ to stabilize the open-loop system (7.1) and regulate the performance $\mathbf{z}$ to zero when the exogenous signal $\mathbf{w}$ is composed of signals generated by an unforced linear system [99]. In this section, the approach taken will be that of the internal model controller described in [99]. This method involves augmenting the plant (7.42) with a model of the disturbance. A feedback controller is then designed to stabilize the augmented plant. The final filter is formed by augmenting the resulting feedback controller with the internal model.

As before, the exogenous signal is given by $\mathbf{w} = \begin{bmatrix} \delta\mathbf{d}^{\mathsf{T}} & \delta\mathbf{n}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$, where, as in (6.31), $\delta\mathbf{d}$ is the output of the linear system

$$\dot{\mathbf{x}}_d = \mathbf{A}_d\mathbf{x}_d, \tag{7.52}$$

$$\delta\mathbf{d} = \mathbf{C}_d\mathbf{x}_d. \tag{7.53}$$

The internal model of the disturbance is selected as

$$\boldsymbol{\mathcal{H}}_{im} \left\{ \begin{array}{ll} \dot{\hat{\mathbf{x}}}_d & = \mathbf{A}_d \hat{\mathbf{x}}_d + \mathbf{B}_d \mathbf{y}, \\ \mathbf{y}_{im} & = \hat{\mathbf{x}}_d + \mathbf{D}_d \mathbf{w}, \end{array} \right. \tag{7.54}$$

where the output $\mathbf{y}_{im}$ of the system $\boldsymbol{\mathcal{H}}_{im}$ is the disturbance state $\mathbf{x}_d$, and $\mathbf{D}_d$ is a design variable.

The matrix $\mathbf{B}_d$ is chosen such that $(\mathbf{A}_d, \mathbf{B}_d)$ is controllable [99]. The plant (7.42) is then augmented with the internal model (7.54) such that the augmented system is given by

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\hat{\mathbf{x}}}_d \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B}_d \mathbf{C}_2 & \mathbf{A}_d \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \hat{\mathbf{x}}_d \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_d \mathbf{D}_{21} \end{bmatrix} \mathbf{w} + \begin{bmatrix} \mathbf{B}_2 \\ \mathbf{B}_d \mathbf{D}_{22} \end{bmatrix} \mathbf{u}, \tag{7.55}$$

$$\mathbf{z} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} \end{bmatrix} \mathbf{x}_a + \mathbf{D}_{11} \mathbf{w} + \mathbf{D}_{12} \mathbf{u}, \tag{7.56}$$

$$\begin{bmatrix} \mathbf{y} \\ \hat{\mathbf{x}}_d \end{bmatrix} = \begin{bmatrix} \mathbf{C}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \hat{\mathbf{x}}_d \end{bmatrix} + \begin{bmatrix} \mathbf{D}_{21} \\ \mathbf{D}_d \end{bmatrix} \mathbf{w} + \begin{bmatrix} \mathbf{D}_{22} \\ \mathbf{0} \end{bmatrix} \mathbf{u}. \tag{7.57}$$

Defining $\mathbf{x}_a = \begin{bmatrix} \mathbf{x}^\mathsf{T} & \hat{\mathbf{x}}_d^\mathsf{T} \end{bmatrix}^\mathsf{T}$, the augmented system can be written as

$$\dot{\mathbf{x}}_a = \mathbf{A}_a \mathbf{x}_a + \mathbf{B}_{a1} \mathbf{w} + \mathbf{B}_{a2} \mathbf{u}, \tag{7.58a}$$

$$\mathbf{z}_a = \mathbf{C}_{a1} \mathbf{x}_a + \mathbf{D}_{a11} \mathbf{w} + \mathbf{D}_{a12} \mathbf{u}, \tag{7.58b}$$

$$\mathbf{y}_a = \mathbf{C}_{a2} \mathbf{x}_a + \mathbf{D}_{a21} \mathbf{w} + \mathbf{D}_{a22} \mathbf{u}, \tag{7.58c}$$

where $\mathbf{y}_a = \begin{bmatrix} \mathbf{y}^\mathsf{T} & \hat{\mathbf{x}}_d^\mathsf{T} \end{bmatrix}^\mathsf{T}$ and $\mathbf{z}_a = \mathbf{z}$ [99]. A feedback controller is now designed to stabilize the augmented plant (7.58). The controller takes the form

$$\boldsymbol{\mathcal{H}}_c \left\{ \begin{array}{ll} \dot{\mathbf{x}}_c & = \mathbf{A}_c \mathbf{x}_c + \mathbf{B}_c \mathbf{y}_a, \\ \mathbf{u} & = \mathbf{C}_c \mathbf{x}_c, \end{array} \right. \tag{7.59}$$

where the transfer function associated with the system $\boldsymbol{\mathcal{H}}_c$ is given by $\mathbf{H}_c(s) = \mathbf{C}_c(s\mathbf{1} - \mathbf{A}_c)^{-1}\mathbf{B}_c$. The feedback controller $\mathbf{H}_c(s)$ is taken as the $\mathcal{H}_2$-optimal filter for the augmented plant (7.58), that is $\mathbf{H}_c(s) = \mathbf{H}_{\mathcal{H}_2}(s)$ defined in (7.43). Specifically,

$$\mathbf{A}_c = \mathbf{A}_a - \mathbf{B}_{a2}\mathbf{K} - \mathbf{L}\mathbf{C}_{a2}, \tag{7.60}$$

$$\mathbf{B}_c = \mathbf{L}, \tag{7.61}$$

$$\mathbf{C}_c = \mathbf{K}, \tag{7.62}$$

where $\mathbf{K} = (\mathbf{D}_{a12}^{\mathsf{T}}\mathbf{D}_{a12})^{-1}\mathbf{B}_{a2}^{\mathsf{T}}\mathbf{P}_a$, $\mathbf{L} = \mathbf{\Pi}_a\mathbf{C}_{a2}^{\mathsf{T}}(\mathbf{D}_{a21}\mathbf{D}_{a21}^{\mathsf{T}})^{-1}$, and the matrices $\mathbf{P}_a = \mathbf{P}_a^{\mathsf{T}} > 0$ and $\mathbf{\Pi}_a = \mathbf{\Pi}_a^{\mathsf{T}} > 0$ are the solutions to the algebraic Riccati equations

$$\begin{aligned} \mathbf{P}_a\mathbf{A}_a + \mathbf{A}_a^{\mathsf{T}}\mathbf{P}_a - \mathbf{P}_a\mathbf{B}_{a2}(\mathbf{D}_{a12}^{\mathsf{T}}\mathbf{D}_{a12})^{-1}\mathbf{B}_{a2}^{\mathsf{T}}\mathbf{P}_a + \mathbf{C}_{a1}^{\mathsf{T}}\mathbf{C}_{a1} &= \mathbf{0}, \qquad (7.63) \\ \mathbf{\Pi}_a\mathbf{A}_a^{\mathsf{T}} + \mathbf{A}_a\mathbf{\Pi}_a - \mathbf{\Pi}_a\mathbf{C}_{a2}^{\mathsf{T}}(\mathbf{D}_{a21}\mathbf{D}_{a21}^{\mathsf{T}})^{-1}\mathbf{C}_{a2}\mathbf{\Pi}_a + \mathbf{B}_{a1}\mathbf{B}_{a1}^{\mathsf{T}} &= \mathbf{0}. \qquad (7.64) \end{aligned}$$

An intermediate filter $\bar{\mathbf{H}}(s) = \bar{\mathbf{C}}_f(s\mathbf{1} - \mathbf{A}_f)^{-1}\mathbf{B}_f$ is formed by augmenting the feedback controller with the internal model such that

$$\mathbf{A}_f = \begin{bmatrix} \mathbf{A}_c & \mathbf{B}_{cd} \\ \mathbf{0} & \mathbf{A}_d \end{bmatrix}, \quad \mathbf{B}_f = \begin{bmatrix} \mathbf{B}_{cy} \\ \mathbf{B}_w \end{bmatrix}, \quad \bar{\mathbf{C}}_f = \begin{bmatrix} \mathbf{C}_c & \mathbf{0} \end{bmatrix}, \qquad (7.65)$$

where the matrix $\mathbf{B}_c$ is partitioned as $\mathbf{B}_c = \begin{bmatrix} \mathbf{B}_{cy} & \mathbf{B}_{cd} \end{bmatrix}$ such that $\mathbf{B}_c\mathbf{y}_a = \mathbf{B}_{cy}\mathbf{y} + \mathbf{B}_{cd}\hat{\mathbf{x}}_d$ [99]. If no SPR restriction is made on the filter in (6.12) then $\bar{\mathbf{H}}(s)$ may be implemented. However, to ensure asymptotic convergence the filter $\bar{\mathbf{H}}(s)$ is modified in the same manner as in Sec. 7.3.4. This is done by solving the optimization problem in (7.51) for the matrix $\mathbf{C}_f$ such that $\mathbf{H}(s) = \mathbf{C}_f(s\mathbf{1} - \mathbf{A}_f)^{-1}\mathbf{B}_f$ is SPR, and $\mathbf{H}(s)$ and $\bar{\mathbf{H}}(s)$ remain as close as possible. Specifically, the optimization problem given by

$$\begin{aligned} \min_{\mathbf{P}_f, \mathbf{Z}, g} \quad & \mathcal{J} = g \qquad (7.66a) \\ \text{subject to} \quad & \mathbf{P}_f = \mathbf{P}_f^{\mathsf{T}} > 0, \qquad (7.66b) \\ & \mathbf{P}_f\mathbf{A}_f + \mathbf{A}_f^{\mathsf{T}}\mathbf{P}_f < 0, \qquad (7.66c) \\ & \mathbf{Z} = \mathbf{Z}^{\mathsf{T}} \geq 0, \qquad (7.66d) \\ & \mathrm{tr}(\mathbf{Z}) \leq g, \qquad (7.66e) \\ & \begin{bmatrix} \mathbf{Z} & \mathbf{B}_f^{\mathsf{T}}\mathbf{P}_f - \bar{\mathbf{C}}_f \\ (\mathbf{B}_f^{\mathsf{T}}\mathbf{P}_f - \bar{\mathbf{C}}_f)^{\mathsf{T}} & \mathbf{1} \end{bmatrix} \geq 0, \qquad (7.66f) \end{aligned}$$

is solved for $\mathbf{P}_f$ and $\mathbf{C}_f$ is set to $\mathbf{C}_f = \mathbf{B}_f^{\mathsf{T}}\mathbf{P}_f$.

The resulting closed-loop system formed by the feedback interconnection between $\boldsymbol{\mathcal{H}}$ and the nonlinear error dynamics (6.18) is shown in Fig. 7.6

7.4.2 Summary

To summarize this section, a review of the synthesis procedure starting from the left observer error dynamics in (6.18) is provided.

1. Linearize the nonlinear system (6.18) as in Proposition 7.3.

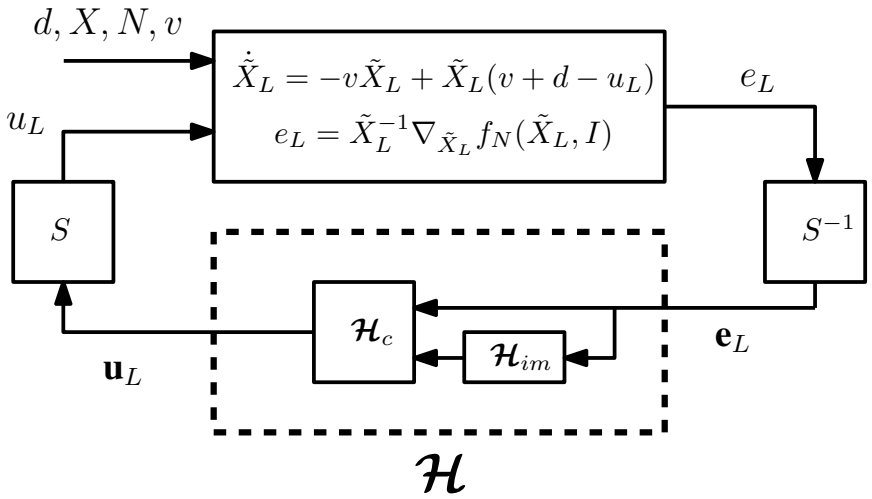


Figure 7.6: Block diagram of the left observer error dynamics with an internal model.

2. Pick weights $\mathbf{W}_d(s)$, $\mathbf{W}_n(s)$, $\mathbf{W}_z(s)$.

3. Formulate the standard control problem in (7.42).

4. Pick $\mathbf{B}_d$ in (7.54) such that $(\mathbf{A}_d, \mathbf{B}_d)$ is controllable.

5. Augment the plant with the internal model (7.54).

6. Find the optimal $\mathcal{H}_2$ feedback controller that stabilizes (7.58).

7. Augment the feedback controller (7.59) with the internal model (7.54).

8. Solve the optimization problem in (7.66) to construct the SPR system $\boldsymbol{\mathcal{H}}$.

7.5 Numerical Example

In this section, the tools developed in the current chapter and in Chapter 6, are applied to the attitude estimation problem previously considered in Chapter 4. Recall that the attitude of a rigid body can be uniquely described by a rotation matrix $\mathbf{R} \in SO(3)$, which evolves according to Poisson's equation (4.4), repeated here as

$$\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\omega}. \tag{7.67}$$

Similar to Chapter 4, it is assumed that angular velocity measurements are available as

$$\boldsymbol{\omega}_y = \boldsymbol{\omega} + \mathbf{d}, \tag{7.68}$$

where $\mathbf{d} \in \mathbb{R}^3$ is some disturbance associated with the angular velocity measurement. It is also assumed that there exists two reference vectors $\mathbf{y}_a^1, \mathbf{y}_a^2 \in \mathbb{S}^2$ available for measurement as

$$\mathbf{y}_b^j = \mathbf{R}^\mathsf{T} \mathbf{N}_j^\mathsf{T} \mathbf{y}_a^j, \ j \in \{1, 2\}, \tag{7.69}$$

where $\mathbf{y}_b^j \in \mathbb{S}^2$ is the measurement of $\mathbf{y}_a^j$ and $\mathbf{N}_j \in SO(3)$ is a group element that introduces noise into the measurement of $\mathbf{y}_a^j$.

A basis for $\mathfrak{so}(3)$ is taken as $B = \{\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3\}$, where $\mathbf{B}_i = \mathbf{e}_i^\times$ and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis of $\mathbb{R}^3$. In this basis, the mapping $S : \mathbb{R}^3 \to \mathfrak{so}(3)$ is given by the operator $(\cdot)^\times$.

7.5.1 Error Function

As in Chapter 4, let $\hat{\mathbf{R}} \in SO(3)$ denote the estimate of $\mathbf{R}$. As a cost function on $SO(3)$, take (4.16), repeated here in terms of vector measurements as

$$\gamma_y(\hat{\mathbf{R}}, y) = \sum_{j=1}^{n} ||\hat{\mathbf{R}}^\mathsf{T} \mathbf{y}_a^j - \mathbf{y}_b^j||_2^2, \tag{7.70}$$

where $y = (\mathbf{y}_a^1, \mathbf{y}_a^2) \in \mathbb{S}^2 \times \mathbb{S}^2$, and $\gamma_y : SO(3) \times \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}^+$. Letting $N = (\mathbf{N}_1, \mathbf{N}_2) \in SO(3) \times SO(3)$, the cost function may be written as

$$\gamma(\hat{\mathbf{R}}, \mathbf{R}, N) = \gamma_N(\hat{\mathbf{R}}, \mathbf{R}) = \sum_{j=1}^{n} ||\hat{\mathbf{R}}^\mathsf{T} \mathbf{y}_a^j - \mathbf{R}^\mathsf{T} \mathbf{N}_j^\mathsf{T} \mathbf{y}_a^j||_2^2. \tag{7.71}$$

The cost function γ_N is right-invariant and $\gamma_\mathcal{I}$ is an error function satisfying all of the conditions of Definition 6.3 [17]. From Proposition 4.1 and (4.99), the gradient of γ_N with respect to $\hat{\mathbf{R}}$ is

$$\boldsymbol{\nabla}_{\hat{\mathbf{R}}} \gamma_N(\hat{\mathbf{R}}, \mathbf{R}) = \mathbf{e}_R^\times \hat{\mathbf{R}}, \tag{7.72}$$

where

$$\mathbf{e}_R = \tfrac{1}{2} \sum_{j=1}^{2} (\mathbf{y}_a^{j\times} \hat{\mathbf{R}} \mathbf{y}_b^j) = \boldsymbol{\mathcal{P}}_a \left(\tilde{\mathbf{R}}_R \sum_{j=1}^{2} \mathbf{N}_j^\mathsf{T} \mathbf{y}_a^j \mathbf{y}_a^{j\mathsf{T}} \right)^\mathsf{V}. \tag{7.73}$$

7.5.2 $SO(3)$ Observer Design

The right observer, given in (6.8), for the Lie group $SO(3)$ is given by

$$\begin{aligned}
\dot{\hat{\mathbf{R}}} &= \hat{\mathbf{R}}\boldsymbol{\omega}_y^\times - \mathbf{u}_R^\times \hat{\mathbf{R}}, && (7.74a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && (7.74b)\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && (7.74c)
\end{aligned}$$

while the left observer, given in (6.12), is

$$\begin{aligned}
\dot{\hat{\mathbf{R}}} &= \hat{\mathbf{R}}\boldsymbol{\omega}_y^\times - \hat{\mathbf{R}}\mathbf{u}_L^\times, && (7.75a)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, && (7.75b)\\
\mathbf{u}_L &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, && (7.75c)
\end{aligned}$$

Recall that $\mathbf{e}_L^\times = \mathrm{Ad}_{\hat{\mathbf{R}}^\mathsf{T}}(\mathbf{e}_R^\times) = \hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_R^\times\hat{\mathbf{R}} = (\hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_R)^\times$ [66]. Consequently, $\mathbf{e}_L = \hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_R$.

7.5.2.1 Linearization

The right and left group errors are defined as $\tilde{\mathbf{R}}_R = \hat{\mathbf{R}}\mathbf{R}^\mathsf{T}$ and $\tilde{\mathbf{R}}_L = \mathbf{R}^\mathsf{T}\hat{\mathbf{R}}$. Both the left and right observers will be implemented and consequently it will be useful to determine the linearizations of their associated error dynamics.

Right Linearization: From Proposition 6.1 the error dynamics associated with $\tilde{\mathbf{R}}_R$ are given by

$$\begin{aligned}
\dot{\tilde{\mathbf{R}}}_R &= \tilde{\mathbf{R}}_R(\mathbf{R}\mathbf{d})^\times - \mathbf{u}_R^\times\tilde{\mathbf{R}}_R && (7.76)\\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, && (7.77)\\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, && (7.78)\\
\mathbf{e}_R &= \mathcal{P}_a\left(\tilde{\mathbf{R}}_R\sum_{j=1}^{2}\mathbf{N}_j\mathbf{y}_a^j\mathbf{y}_a^{j\mathsf{T}}\right)^\mathsf{V}. && (7.79)
\end{aligned}$$

Let $\tilde{\mathbf{R}}_R = \exp(\tilde{\boldsymbol{\phi}}_R^\times)$, where $\tilde{\boldsymbol{\phi}}_R \in \mathbb{R}^3$. Further, let $\mathbf{N}_j = \exp(\mathbf{n}_j^\times)$, where $\mathbf{n}_j \in \mathbb{R}^3$, $j \in \{1,2\}$. Define $\mathbf{n} = \begin{bmatrix} \mathbf{n}_1^\mathsf{T} & \mathbf{n}_2^\mathsf{T} \end{bmatrix}^\mathsf{T}$. The rotation matrix $\tilde{\mathbf{R}}_R$ can be perturbed about $\mathbf{1}$ by letting $\boldsymbol{\phi}_R = \bar{\tilde{\boldsymbol{\phi}}}_R + \delta\tilde{\boldsymbol{\phi}}_R$, where $\bar{\tilde{\boldsymbol{\phi}}}_R = \mathbf{0}$ is the nominal value of $\tilde{\boldsymbol{\phi}}_R$. Similarly, let $\mathbf{n}_j = \bar{\mathbf{n}}_j + \delta\mathbf{n}_j$ and $\mathbf{d} = \bar{\mathbf{d}} + \delta\mathbf{d}$, where $\bar{\mathbf{n}} = \mathbf{0}$ and $\bar{\mathbf{d}} = \mathbf{0}$. Then, from Proposition 7.2,

the linearization of the right error dynamics are given by

$$\delta\dot{\tilde{\boldsymbol{\phi}}}_R = \delta\mathbf{d}' - \delta\mathbf{u}_R, \tag{7.80a}$$

$$\dot{\mathbf{x}}_f = \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f(\mathbf{M}_1\delta\tilde{\boldsymbol{\phi}}_R + \mathbf{M}_2\delta\mathbf{n}), \tag{7.80b}$$

where $\delta\mathbf{d}' = \hat{\mathbf{R}}\delta\mathbf{d}$. To find expressions for $\mathbf{M}_1$ and $\mathbf{M}_2$ consider

$$\mathbf{e}_R^\times = \boldsymbol{\mathcal{P}}_a\left(\tilde{\mathbf{R}}_R\sum_{j=1}^{2}\mathbf{N}_j^\mathsf{T}\mathbf{y}_a^j\mathbf{y}_a^{j\mathsf{T}}\right) \tag{7.81}$$

$$= \boldsymbol{\mathcal{P}}_a\left((\mathbf{1}+\delta\tilde{\boldsymbol{\phi}}_R^\times)\sum_{j=1}^{2}(\mathbf{1}-\delta\mathbf{n}_j^\times)\mathbf{y}_a^j\mathbf{y}_a^{j\mathsf{T}}\right), \tag{7.82}$$

where the small angle approximations $\tilde{\mathbf{R}}_R \approx \mathbf{1}+\delta\tilde{\boldsymbol{\phi}}_R^\times$ and $\mathbf{N}_j \approx \mathbf{1}+\delta\mathbf{n}_j^\times$ have been used [66]. Neglecting higher order terms in $\delta\tilde{\boldsymbol{\phi}}_R$ and $\delta\mathbf{n}_j$ gives

$$\delta\mathbf{e}_R^\times = \boldsymbol{\mathcal{P}}_a\left(\delta\tilde{\boldsymbol{\phi}}_R^\times\mathbf{M}_a\right) - \boldsymbol{\mathcal{P}}_a\left(\sum_{j=1}^{2}\mathbf{n}_j^\times\mathbf{y}_a^j\mathbf{y}_a^{j\mathsf{T}}\right) \tag{7.83}$$

$$= \tfrac{1}{2}([\mathrm{tr}(\mathbf{M}_a)\mathbf{1}-\mathbf{M}_a]\delta\tilde{\boldsymbol{\phi}}_R)^\times + \tfrac{1}{2}\sum_{j=1}^{2}(\mathbf{y}_a^{j\times}\mathbf{y}_a^{j\times}\delta\mathbf{n}_j)^\times, \tag{7.84}$$

which follows from the identities in (4.6) and (4.7). This implies that

$$\delta\mathbf{e}_R = \mathbf{M}_1\delta\tilde{\boldsymbol{\phi}}_R + \mathbf{M}_2\delta\mathbf{n}. \tag{7.85}$$

where

$$\mathbf{M}_1 = \tfrac{1}{2}(\mathrm{tr}(\mathbf{M}_a)\mathbf{1}-\mathbf{M}_a), \text{ and, } \mathbf{M}_2 = \begin{bmatrix} \tfrac{1}{2}\mathbf{y}_a^{1\times}\mathbf{y}_a^{1\times} & \tfrac{1}{2}\mathbf{y}_a^{2\times}\mathbf{y}_a^{2\times} \end{bmatrix}. \tag{7.86}$$

Left Linearization: The error function γ_N is not bi-invariant and consequently the error dynamics in Proposition 6.2 do not hold in general. However, while (6.16d) does not hold, it can be shown that Equations (6.16a)-(6.16c) are valid and consequently,

$$\dot{\tilde{\mathbf{R}}}_L = -\boldsymbol{\omega}^\times\tilde{\mathbf{R}}_L + \tilde{\mathbf{R}}_L(\boldsymbol{\omega}+\mathbf{d}-\mathbf{u}_L) \tag{7.87}$$

$$\dot{\mathbf{x}}_f = \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_L, \tag{7.88}$$

$$\mathbf{u}_L = \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_L, \tag{7.89}$$

$$\mathbf{e}_L = \boldsymbol{\mathcal{P}}_a\left(\mathbf{R}^\mathsf{T}\sum_{j=1}^{2}\mathbf{N}_j^\mathsf{T}\mathbf{y}_a^j\mathbf{y}_a^{j\mathsf{T}}\hat{\mathbf{R}}\right)^\vee, \tag{7.90}$$

where the relation $\mathbf{e}_L^\times = \hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_R^\times\hat{\mathbf{R}}$ has been used. Let $\tilde{\mathbf{R}}_L = \exp(\tilde{\boldsymbol{\phi}}_L^\times)$, where $\tilde{\boldsymbol{\phi}}_L \in \mathbb{R}^3$ and consider a perturbation about $\tilde{\mathbf{R}}_L = \mathbf{1}$ by letting $\tilde{\boldsymbol{\phi}}_L = \bar{\tilde{\boldsymbol{\phi}}}_L + \delta\tilde{\boldsymbol{\phi}}_L$ with $\bar{\tilde{\boldsymbol{\phi}}}_L = \mathbf{0}$. With a small angle assumption, $\tilde{\mathbf{R}}_L \approx \mathbf{1} + \delta\tilde{\boldsymbol{\phi}}_L^\times$. Again, as γ_N is not bi-invariant the linearization given in Proposition 7.3 does not hold completely. However, it can be shown that (7.25a) does hold and

$$\delta\dot{\tilde{\boldsymbol{\phi}}}_L = \mathbf{A}(t)\delta\tilde{\boldsymbol{\phi}}_L + \delta\mathbf{d} - \delta\mathbf{u}_L. \tag{7.91}$$

Recall $\mathbf{A}(t) = [-\mathrm{ad}_v]_B$, where v was the group velocity. In the case of the Lie group $SO(3)$ the group velocity is given by $\boldsymbol{\omega}^\times$ and consequently from (4.5) and (4.6),

$$\begin{aligned} -\mathrm{ad}_{\boldsymbol{\omega}^\times}(\delta\tilde{\boldsymbol{\phi}}_L^\times) &= -[\boldsymbol{\omega}^\times, \delta\tilde{\boldsymbol{\phi}}_L^\times] && (7.92)\\ &= -\boldsymbol{\omega}^\times\delta\tilde{\boldsymbol{\phi}}_L^\times + \delta\tilde{\boldsymbol{\phi}}_L^\times\boldsymbol{\omega}^\times && (7.93)\\ &= -(\boldsymbol{\omega}^\times\delta\tilde{\boldsymbol{\phi}})^\times. && (7.94)\end{aligned}$$

Therefore $[-\mathrm{ad}_v]_B = -\boldsymbol{\omega}^\times$ and

$$\delta\dot{\tilde{\boldsymbol{\phi}}}_L = -\boldsymbol{\omega}^\times\delta\tilde{\boldsymbol{\phi}}_L + \delta\mathbf{d} - \delta\mathbf{u}_L. \tag{7.95}$$

To find an expression for $\delta\mathbf{e}_L$, recall that $\mathbf{e}_L = \hat{\mathbf{R}}^\mathsf{T}\mathbf{e}_R$ and consequently $\delta\mathbf{e}_L = \hat{\mathbf{R}}^\mathsf{T}\delta\mathbf{e}_R$. Further $\tilde{\mathbf{R}}_L = \hat{\mathbf{R}}^\mathsf{T}\tilde{\mathbf{R}}_R\hat{\mathbf{R}}$ and thus perturbing both $\tilde{\mathbf{R}}_L$ and $\tilde{\mathbf{R}}_R$ about $\mathbf{1}$ gives

$$\begin{aligned} \mathbf{1} + \delta\tilde{\boldsymbol{\phi}}_L^\times &= \hat{\mathbf{R}}^\mathsf{T}(\mathbf{1} + \delta\tilde{\boldsymbol{\phi}}_R^\times)\hat{\mathbf{R}}\\ \delta\tilde{\boldsymbol{\phi}}_L^\times &= \hat{\mathbf{R}}^\mathsf{T}\delta\tilde{\boldsymbol{\phi}}_R^\times\mathbf{R},\end{aligned}$$

and thus $\delta\tilde{\boldsymbol{\phi}}_L = \hat{\mathbf{R}}^\mathsf{T}\tilde{\boldsymbol{\phi}}_R$. Consequently, as $\delta\mathbf{e}_R = \mathbf{M}_1\delta\tilde{\boldsymbol{\phi}}_R + \mathbf{M}_2\delta\mathbf{n}$, it follows that

$$\delta\mathbf{e}_L = \hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_1\hat{\mathbf{R}}\delta\tilde{\boldsymbol{\phi}}_L + \hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_2\delta\mathbf{n}. \tag{7.96}$$

The lack of symmetry between $\hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_1\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_2$ is due to the fact that noise is introduced into the vector measurements by left translating $\mathbf{R}$ by $\mathbf{N}_j$, $j \in \{1, 2\}$, such that $\mathbf{y}_b^j = \mathbf{R}^\mathsf{T}\mathbf{N}_j^\mathsf{T}\mathbf{y}_a^j$. Had noise been introduced by right translation such that $\mathbf{y}_b^j = \mathbf{N}_j^\mathsf{T}\mathbf{R}^\mathsf{T}\mathbf{y}_a^j$ then it could be shown that $\delta\mathbf{e}_L = \hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_1\hat{\mathbf{R}}\delta\tilde{\boldsymbol{\phi}}_L + \hat{\mathbf{R}}^\mathsf{T}\mathbf{M}_2\hat{\mathbf{R}}\delta\mathbf{n}$.

7.5.3 Simulation Parameters

The angular velocity of the rigid body is described by

$$\boldsymbol{\omega}(t) = -\pi^2 \cos(\pi/10t)/60.0 \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathsf{T}} \text{ (rad/s)}. \tag{7.97}$$

The initial attitude is set to $\mathbf{R}(0) = \exp(\boldsymbol{\phi}(0)^{\times})$, where

$$\boldsymbol{\phi}(0) = \pi/6 \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathsf{T}} \text{ (rad)}, \tag{7.98}$$

and the two reference vectors are given by $\mathbf{y}_a^1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathsf{T}}$ and $\mathbf{y}_a^2 = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^{\mathsf{T}}$. For the following simulations, the observers are initialized with $\hat{\mathbf{R}}(0) = \mathbf{1}$ and $\mathbf{x}_f(0) = \mathbf{0}$. The multiplicative noise associated with the reference measurements, $\mathbf{N}_j = \exp(\mathbf{n}_j^{\times})$, are constructed by selecting $\mathbf{n}_j$ as linear combinations of harmonic signals with frequencies in the range $[20, 48]$ (rad/s). Specifically,

$$\mathbf{n}_1(t) = \begin{bmatrix} 0.025\sin(20t + 0.5) + 0.15\cos(32t + 0.1) \\ 0.1\sin(35t + 0.4) + 0.08\cos(20t + 0.5) \\ 0.2\sin(48t + 0.2) + 0.18\cos(40t + 0.9\pi) \end{bmatrix} \text{ (rad)}, \tag{7.99}$$

and

$$\mathbf{n}_2(t) = \begin{bmatrix} 0.15\sin(24t + 0.6) + 0.05\cos(42t + 0.5) \\ 0.03\sin(42t + 1.5) + 0.25\cos(34t + 0.2) \\ 0.1\sin(36t + 0.2) + 0.15\cos(27t + 1.2\pi) \end{bmatrix} \text{ (rad)}. \tag{7.100}$$

7.5.4 Linear Filter Design

Two separate methodologies are considered for the design of the filter $\mathbf{H}(s)$. In the first, $\mathbf{H}(s)$ is designed by first taking the Laplace transform of the linearized system. Then, classical control techniques can be used to shape the sensitivity and closed-loop sensitivity transfer functions based on the frequency content of the exogenous signals. In the second method, the optimal observer design strategy developed in this chapter is used. In this case, the filter $\mathbf{H}(s)$ is designed based on minimizing the $\mathcal{H}_2$-norm of the closed-loop system.

Method 1: To simplify the design of the filter, let $\mathbf{H}(s) = \mathbf{C}_f(s\mathbf{1} - \mathbf{A}_f)^{-1}\mathbf{B}_f + \mathbf{D}_f = H(s)\mathbf{M}_1^{-1}$. Then, taking the Laplace transform of (7.80) yields

$$\begin{aligned} \delta\tilde{\boldsymbol{\phi}}_R(s) &= \frac{s}{1+H(s)}\frac{\delta\mathbf{d}'(s)}{s} + \frac{H(s)}{s+H(s)}\mathbf{M}_1^{-1}\mathbf{M}_2\delta\mathbf{n}(s) \\ &= S(s)\frac{\delta\mathbf{d}'(s)}{s} + T(s)\mathbf{M}_1^{-1}\mathbf{M}_2\delta\mathbf{n}(s). \end{aligned}$$

The undesirable effects of the partial state measurement noise can be mitigated by designing $T(s)$ as a low-pass filter with an appropriate cutoff frequency. Consider two versions of $H(s)$,

$$H_1(s) = k, \tag{7.101}$$

and

$$H_2(s) = b/(s+a), \tag{7.102}$$

where $k, a, b \in (0, \infty)$. By selecting $H_1(s)$ the transfer function $T(s)$ is a first order low-pass filter of the form $T(s) = k/(s+k)$. This is equivalent to the attitude observer proposed in [10]. Alternatively, selecting the SPR transfer function $H_2(s)$ results in a second order low-pass filter $T(s) = b/(s^2 + as + b)$. Selecting $k = 5$, $a = 15.5$, and $b = 60.4$ results in a cutoff frequency of 5 (rad/s) for $T(s)$. Although $H_1(s)$ and $H_2(s)$ give the same cutoff frequency for $T(s)$, the second order low-pass filter rolls off at -40 dB per decade while the first order low-pass filter rolls off at -20 dB per decade. Consequently, it is expected that greater noise mitigation can be accomplished by selecting $H_2(s)$ over $H_1(s)$. The linear filters given by $H_1(s)\mathbf{M}_1^{-1}$ and $H_2(s)\mathbf{M}_1^{-1}$ are denoted $\mathbf{H}_1(s)$ and $\mathbf{H}_2(s)$ respectively.

Method 2: The second method considered is the $\mathcal{H}_2$-optimal filter design procedure developed in Sec. 7.3. This involves selecting the weights $\mathbf{W}_d(s)$, $\mathbf{W}_n(s)$, $\mathbf{W}_{z_1}(s)$, and $\mathbf{W}_{z_2}(s)$, as well as the design variable r_1. The weights are taken as $\mathbf{W}_d(s) = \mathbf{W}_{z_1}(s) = 5\pi/(s+5\pi)\mathbf{1}$, $\mathbf{W}_{z_2}(s) = 10^{-4}\cdot\mathbf{1}$, and $\mathbf{W}_n(s) = (10s+45\pi)/(s+5\pi)\mathbf{1}$. The weights $\mathbf{W}_d(s)$ and $\mathbf{W}_n(s)$ are selected as low- and high-pass filters respectively, which has the effect of weighting the low-frequency content of $\mathbf{d}(s)$ and the high-frequency content of $\mathbf{n}(s)$. Following the steps outlined in Sec. 7.3.5 yields the $\mathcal{H}_2$-optimal filter, which is denoted $\mathbf{H}_3(s)$.

7.5.5 Simulation Results

Case 1 (partial state measurement noise): In the first simulation no disturbances are added to the angular velocity measurement, that is $\mathbf{d}(t) = \mathbf{0}$. The three linear filters $\mathbf{H}_1(s)$, $\mathbf{H}_2(s)$, and $\mathbf{H}_3(s)$ are implemented with the right observer given in (7.74). The results of the simulation are shown in Fig. 7.7 As predicted the steady-

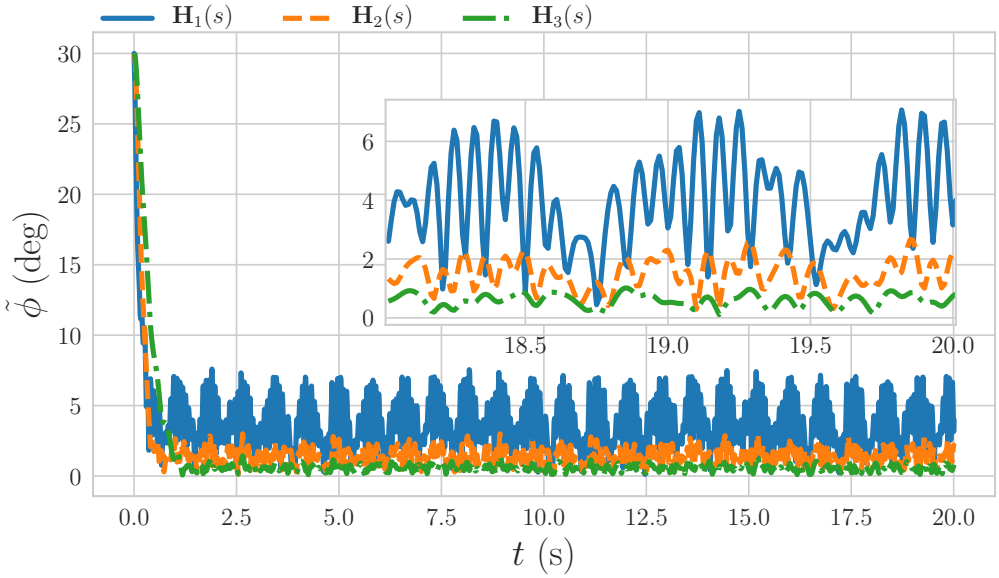


Figure 7.7: Simulation results for case 1. The attitude estimation error for all observers approaches zero.

state error associated with $\mathbf{H}_2(s)$ is indeed lower that that of the constant transfer function, $\mathbf{H}_1(s) = k\mathbf{M}_1^{-1}$. The steady-state error associated with the $\mathcal{H}_2$-optimal filter is also lower than that of both $\mathbf{H}_1(s)$ and $\mathbf{H}_2(s)$.

Case 2 (input disturbance): Suppose now that in addition to partial state measurement noise, the angular velocity measurement is corrupted by a constant bias given by $\mathbf{d}(t) = \begin{bmatrix} 0.1 & -0.1 & 0.1 \end{bmatrix}^{\mathsf{T}}$ (rad/s). Simulation results for the right observer with the constant bias in the angular velocity measurement is shown in Fig. 7.8. The steady-state performance of (7.74) with $\mathbf{H}_1(s)$, $\mathbf{H}_2(s)$, and $\mathbf{H}_3(s)$ observed in the previous simulation has been lost. The poor performance of the observers in this case motivates the introduction of the disturbance observer discussed in Sec. 6.5.

Case 3 (disturbance observer): In an effort to regain the steady-state performance of the observers in the presence of constant bias, the disturbance observer presented in Chapter 6 Section 6.5 is now implemented. The disturbance $\mathbf{d}$ can be written as the output of the linear system in (6.31) with $\mathbf{A}_d = \mathbf{0}$ and $\mathbf{C}_d = \mathbf{1}$ with appropriate initial

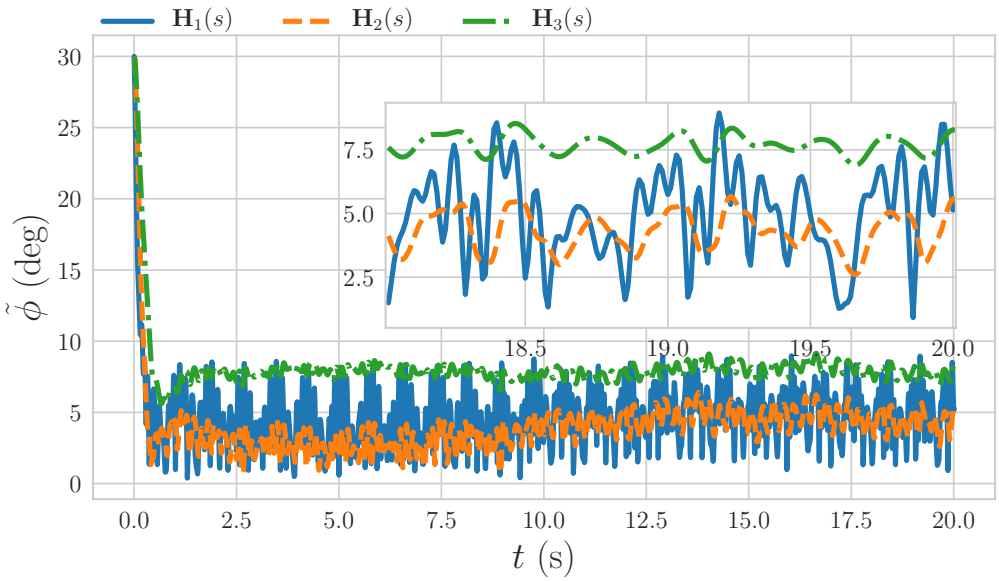


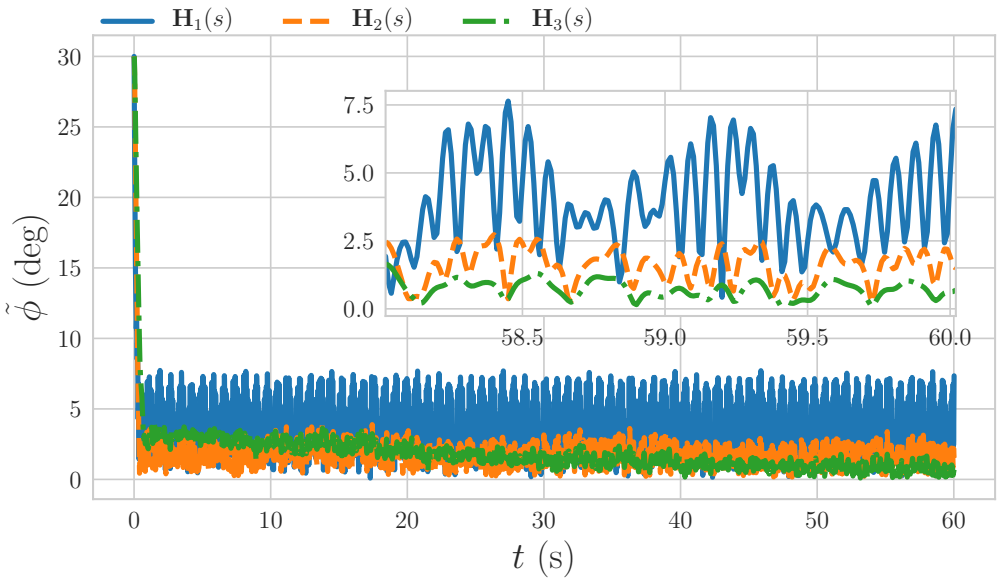
Figure 7.8: Simulation results for case 2. The introduction of the constant bias disturbance causes undesirable effects in the attitude estimate.

conditions. Then, the proposed right observer on $SO(3)$ with associated disturbance observer are given by

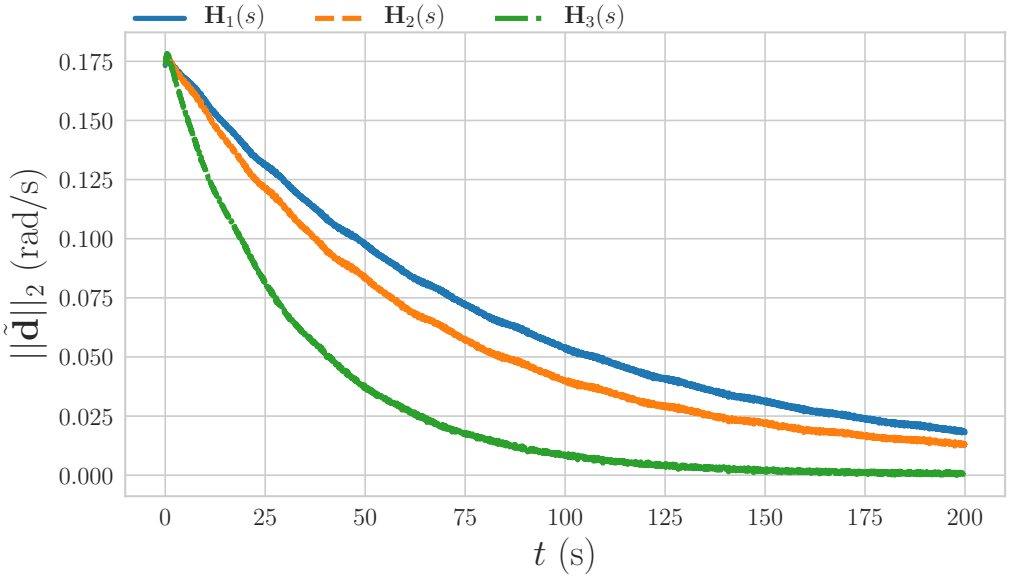
$$\begin{aligned}
\dot{\hat{\mathbf{R}}} &= \hat{\mathbf{R}}\boldsymbol{\omega}_y^{\times} - \hat{\mathbf{R}}\hat{\mathbf{d}}^{\times} - \mathbf{u}_R^{\times}\hat{\mathbf{R}}, \\
\dot{\mathbf{x}}_f &= \mathbf{A}_f\mathbf{x}_f + \mathbf{B}_f\mathbf{e}_R, \\
\mathbf{u}_R &= \mathbf{C}_f\mathbf{x}_f + \mathbf{D}_f\mathbf{e}_R, \\
\dot{\hat{\mathbf{x}}}_d &= \mathbf{A}_d\hat{\mathbf{x}}_d + \rho\mathbf{C}_d^{\mathsf{T}}\bar{\mathbf{e}}_R, \\
\hat{\mathbf{d}} &= \mathbf{C}_d\hat{\mathbf{x}}_d,
\end{aligned}$$

where $\bar{\mathbf{e}}_R = \hat{\mathbf{R}}\mathbf{e}_R$. The right observer on $SO(3)$ with the constant gain transfer function $\mathbf{H}_1(s)$ coupled with a constant bias disturbance observer is of the same form as the observers previously considered in Chapter 4. The results of a simulation with initial conditions $(\hat{\mathbf{R}}, \mathbf{x}_f, \hat{\mathbf{x}}_d) = (\mathbf{1}, \mathbf{0}, \mathbf{0})$ and $\rho = 0.1$ is shown in Fig. 7.9. The attitude estimation error is shown in Fig. 7.9a while the error associated with the bias estimates are shown in Fig. 7.9b. Referring to Fig. 7.9b, it can be seen that the disturbance observer is successful in tracking the true bias as $\tilde{\mathbf{d}}$ approaches zero for all observers. Consequently, the steady-state attitude error observed previously in Fig. 7.7 has been recovered as shown in Fig. 7.9a.

Case 4 (internal model): An alternative method for accomplishing disturbance rejection was discussed in Sec. 7.4. It was assumed in Sec. 7.4 that the error function



(a) Time history of attitude error.



(b) Time history of bias estimation error.

Figure 7.9: Simulation results for case 3.

was bi-invariant, however, the error function currently under consideration, γ_N, does not satisfy this requirement. Due to this fact, the linearized equations associated with the left observer contain time-varying matrices in (7.96) rather than constant matrices as was observed in (7.25b). The filter synthesis procedure in Sec. 7.4 can still be applied by ignoring the time-varying matrices and continuing with the assumption that γ_N is bi-invariant. As will soon be shown, the resulting filter still achieves

satisfactory results, even with the faulty assumption. The same weights used to synthesize $\mathbf{H}_3(s)$ are used in the synthesis of the $\mathcal{H}_2$-optimal filter with an internal model, which is denoted $\mathbf{H}_4(s)$. It remains to select $\mathbf{B}_d$ and $\mathbf{D}_d$ in (7.54). As in the disturbance observer, $\mathbf{B}_d$ is taken as $\mathbf{B}_d = \rho \mathbf{C}_d^\mathsf{T}$ with $\rho = 0.1$. The design variable $\mathbf{D}_d$ is taken as $\mathbf{D}_d = \begin{bmatrix} 100 \cdot \mathbf{1} & \mathbf{0} & \mathbf{0} \end{bmatrix}$. Then, the $\mathcal{H}_2$-optimal filter with internal model, $\mathbf{H}_4(s)$, is synthesized following the steps outlined in Sec. 7.4.2. The results of a simulation with the left observer (7.75) with linear filter $\mathbf{H}_4(s)$ is shown in Fig. 7.10. For comparison purposes, the results from the previous case are also plotted. Referring to Fig. 7.10, the left observer with an internal model is able to successfully reject the constant bias in the angular velocity measurement. The use of $\mathbf{H}_4(s)$ also results in faster convergence to steady-state, compared to the previously considered observers.

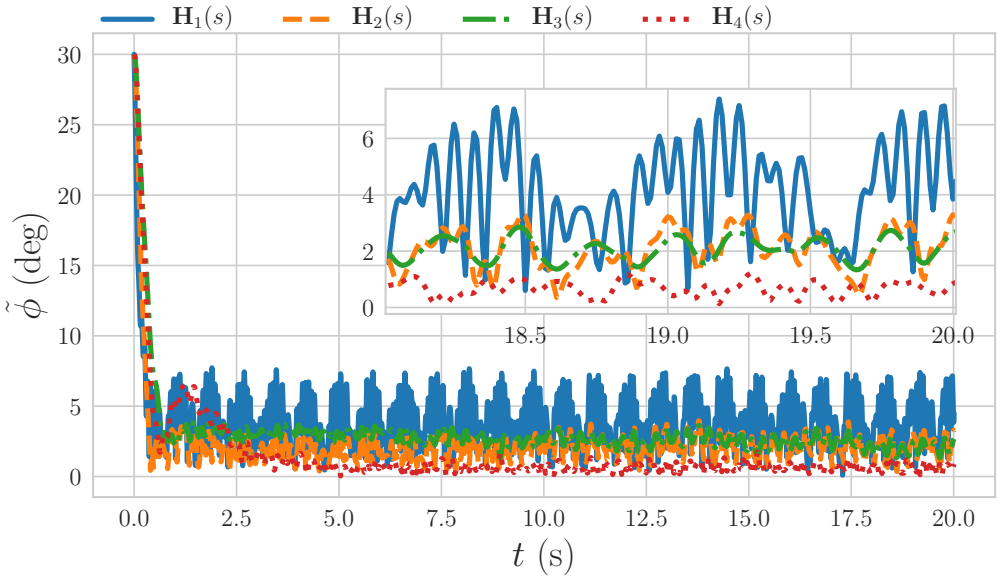


Figure 7.10: Simulation results for case 4. The left observer with an internal model outperforms the other observers.

7.6 Closing Remarks

Optimal design of the linear operator $\boldsymbol{\mathcal{H}}$ for use in the Lie group observers presented in Chapter 6 has been considered in this chapter. Numerical results demonstrated the effectiveness of the proposed observer synthesis methods. Several assumptions and simplifications were made in this chapter. Specifically, time-varying parameters in the linearization of the error dynamics were ignored to simplify the

filter design process. Future work will focus on the application of filter synthesis methods for LTV systems. In addition, only the problem of $\mathcal{H}_2$-optimal estimation was considered. Robust Lie group observer design will be considered in the future by applying $\mathcal{H}_\infty$-optimal observer design strategies.

Part IV

Conclusion

Chapter 8

Closing Remarks and Future Work

This dissertation considered the topic of observer design for systems whose state naturally evolves on a Lie group. Briefly, the main contributions of this work are

1. the development of nonlinear attitude observers with desirable convergence properties,
2. a SLAM algorithm with provable stability guarantees,
3. the design and proof of stability of a higher-order nonlinear complementary filter,
4. $\mathcal{H}_2$-optimal filter synthesis methods, and
5. an alternative method for disturbance rejection in Lie group observers via the internal model principle.

A more complete list of novel contributions of this dissertation can be found in the Preface.

In Part II, two novel attitude estimation algorithms were proposed as well as a provably convergent SLAM algorithm. The SLAM algorithm developed in this dissertation, along with the method presented in [76], are the first steps in the development of a new branch of nonlinear SLAM methods. Several issues remain before the proposed method can be implemented in practice. First, is the problem of data association. As the developed method is deterministic it provides no estimate of the covariance associated with the state estimate. This information is vital in the data association problem which seeks to determine if a measurement corresponds to a previously observed landmark. It may be possible overcome this problem by using uncertainty propagation techniques previously used for the attitude and pose estimation problems. For example, in [113] a deterministic attitude estimator is employed.

The uncertainty in the state estimate is then obtained by propagating an uncertainty ellipsoid along with the state estimates by the use of a Lie group variational integrator. This method may be employed in the context of deterministic SLAM algorithms to recover the state uncertainty necessary for the data association problem. It may also be possible to use the proposed algorithm in tandem with previously developed SLAM methods. For example, due to the computational simplicity of the proposed algorithm many gradient-based observers may be run concurrently. A particle filter may then be implemented where the initial conditions of each gradient-based observer is taken by sampling the current estimate of the PDF.

Part III extended the state of the art in Lie group observer design by the introduction of the higher-order nonlinear complementary filter. Two versions of the higher-order filter were proposed, the right and left observers. Several restrictions were made to the cost function f and to the linear operator $\boldsymbol{\mathcal{H}}$ to ensure that the right and left observers were provably convergent. While the cost function and Riemannian metric were only required to be right-invariant for the right observer they were required to be bi-invariant for the left observer. The bi-invariance condition of f and the Riemannian metric is quite restrictive in that a bi-invariant metric is only guaranteed to exist provided the Lie group is compact. Future work will focus on the relaxing the bi-invariance assumption on f and the Riemannian metric for the case of the left observer. Moreover, the linear operator $\boldsymbol{\mathcal{H}}$ was assumed linear time-invariant as well as SPR. As was seen in Chapter 7, many simplifications were made to ensure that the plant, for which the filter was being designed, was LTI. By relaxing the condition that the linear operator by time-invariant it may be possible to achieve better results by application of linear parameter-varying control techniques or other LTV control design methods. The SPR condition on $\boldsymbol{\mathcal{H}}$ may also be unnecessarily conservative. Recent work in conic-sector-based controller synthesis has the potential to relax the SPR constraint while greatly expanding the types of filters that may be employed [114, 115]. Another area of potential future research is the development of robust observer design on Lie groups by the development of an $\mathcal{H}_\infty$-optimal synthesis method. This may allow for the implementation of robust observer design strategies previously implemented for systems evolving on vector spaces, namely [34]. By posing the filter design problem as an $\mathcal{H}_\infty$ controller design problem, uncertainties in the exogenous signals can be explicitly accounted for.

In summary, this dissertation has explored several applications and novel methods for observer design on Lie groups. Many of the methods presented in this dissertation are in their infancy and require further work before they may be tractable in practice.

However, the theoretical foundations have been laid and it is hoped that the theory outlined in this dissertation will serve as the base of future state estimation techniques.

Bibliography

Bibliography

[1] D. E. Zlotnik and J. R. Forbes, "Nonlinear estimator design on the special orthogonal group using vector measurements directly," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 149–160, 2017.

[2] D. E. Zlotnik and J. R. Forbes, "A nonlinear attitude estimator with desirable convergence properties," in *European Control Conference (ECC)*, pp. 2103–2107, IEEE, 2015.

[3] D. E. Zlotnik and J. R. Forbes, "Exponential convergence of a nonlinear attitude estimator," *Automatica*, vol. 72, pp. 11–18, 2016.

[4] D. E. Zlotnik and J. R. Forbes, "Gradient-based observer for simultaneous localization and mapping," *IEEE Transactions on Automatic Control*, 2018.

[5] D. E. Zlotnik and J. R. Forbes, "Higher-order nonlinear complementary filtering on lie groups," *IEEE Transactions on Automatic Control*, pp. 1–1, 2018.

[6] D. E. Zlotnik and J. R. Forbes, "Exteroceptive measurement filtering embedded within an SO(3)-based attitude estimator," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 296–301, IEEE, 2016.

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[8] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, *et al.*, "Stanley: The robot that won the darpa grand challenge," *Journal of field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[9] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT press, 2005.

[10] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[11] S. Bonnabel, P. Martin, and P. Rouchon, "Symmetry-preserving observers," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2514–2526, 2008.

[12] C. Lageman, R. Mahony, J. Trumpf, *et al.*, "State observers for invariant dynamics on a Lie group," in *18th International Symposium on Mathematical Theory of Networks and Systems*, p. 8, 2008.

[13] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetry-preserving observers on Lie groups," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1709–1713, 2009.

[14] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 367–377, 2010.

[15] A. Khosravian, J. Trumpf, R. Mahony, and C. Lageman, "Bias estimation for invariant systems on Lie groups with homogeneous outputs," in *IEEE Conference on Decision and Control (CDC)*, pp. 4454–4460, IEEE, 2013.

[16] A. Khosravian, J. Trumpf, R. Mahony, and C. Lageman, "Observers for invariant systems on Lie groups with biased input measurements and homogeneous outputs," *Automatica*, vol. 55, pp. 19–26, 2015.

[17] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems*. New York: Springer Science + Business Media, Inc., 2005.

[18] S. Salcudean, "A globally convergent angular velocity observer for rigid body motion," *IEEE Transactions on Automatic Control*, vol. 36, no. 12, pp. 1493–1497, 1991.

[19] H. Nijmeijer and T. I. Fossen, *New directions in nonlinear observer design*, vol. 244. Springer, 1999.

[20] J. Thienel and R. M. Sanner, "A coupled nonlinear spacecraft attitude controller and observer with an unknown constant gyro bias and gyro noise," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2011–2015, 2003.

[21] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[22] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Complementary filter design on the special orthogonal group $SO(3)$," *Proc. of the IEEE Conference on Decision and Control and European Control Conference, Seville, Spain, December 12-15, 2005, pp. 1477–1484.*

[23] M.-D. Hua, "Attitude estimation for accelerated vehicles using GPS/INS measurements," *Control Engineering Practice*, vol. 18, no. 7, pp. 723–732, 2010.

[24] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi, "Attitude estimation using biased gyro and vector measurements with time-varying reference vectors," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1332–1338, 2012.

[25] A. Khosravian and M. Namvar, "Rigid body attitude control using a single vector measurement and gyro," *IEEE Transactions on Automatic Control*, vol. 57, pp. 1273–1279, May 2012.

[26] M.-D. Hua, M. Zamani, J. Trumpf, R. Mahony, and T. Hamel, "Observer design on the special Euclidean group $SE(3)$," in *IEEE Conference on Decision and Control and European Control Conference*, pp. 8169–8175, IEEE, 2011.

[27] M.-D. Hua, T. Hamel, R. Mahony, and J. Trumpf, "Gradient-like observer design on the special Euclidean group $SE(3)$ with system outputs on the real projective space," in *IEEE Conference on Decision and Control (CDC)*, pp. 2139–2145, Dec 2015.

[28] N. A. Chaturvedi, A. K. Sanyal, and N. H. McClamroch, "Rigid-body attitude control," *IEEE Control Systems*, vol. 31, no. 3, pp. 30–51, 2011.

[29] M. Montemerlo and S. Thrun, *FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*, vol. 27. Springer, 2007.

[30] S. Thrun and M. Montemerlo, "The graph slam algorithm with apsplications to large-scale mapping of urban structures," *The International Journal of Robotics Research*, vol. 25, no. 5-6, pp. 403–429, 2006.

[31] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.

[32] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

[33] J.-K. Park, D.-R. Shin, and T. M. Chung, "Dynamic observers for linear time-invariant systems," *Automatica*, vol. 38, no. 6, pp. 1083 – 1087, 2002.

[34] H. Marquez, "A frequency domain approach to state estimation," *Journal of the Franklin Institute*, vol. 340, no. 2, pp. 147–157, 2003.

[35] H. Marquez and M. Riaz, "Robust state observer design with application to an industrial boiler system," *Control Engineering Practice*, vol. 13, no. 6, pp. 713–728, 2005.

[36] A. M. Pertew, H. J. Marquez, and Q. Zhao, "$\mathcal{H}_\infty$ observer design for lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1211–1216, 2006.

[37] M. R. Davoodi, H. Talebi, and H. R. Momeni, "A novel simultaneous fault detection and control approach based on dynamic observer," in *Proceedings of 18th IFAC World Congress, Italy*, pp. 12036–12041, 2011.

[38] S. Andrilli and D. Hecker, *Elementary linear algebra*. Elsevier, 2003.

[39] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*, vol. 36. Springer Science & Business Media, 2013.

[40] D. S. Bernstein, *Matrix mathematics: theory, facts, and formulas*. Princeton University Press, 2009.

[41] A. Bloch, J. Baillieul, P. Crouch, J. E. Marsden, P. S. Krishnaprasad, R. Murray, and D. Zenkov, *Nonholonomic Mechanics and Control*, vol. 24. Springer, 2003.

[42] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[43] J. M. Lee, *Introduction to Smooth Manifold*. Springer, 2003.

[44] P. Petersen, *Riemannian Geometry*, vol. 171. Springer, 2006.

[45] C. Lageman, J. Trumpf, and R. Mahony, "Observers for systems with invariant outputs," in *Control Conference (ECC), 2009 European*, pp. 4587–4592, IEEE, 2009.

[46] R. Mahony, S.-H. Cha, and T. Hamel, "A coupled estimation and control analysis for attitude stabilisation of mini aerial vehicles," *Proc. of the Australasian Conference on Robotics and Automation, Auckland, New Zealand, December 6-8, 2006*.

[47] T. Hamel and R. Mahony, "Attitude estimation on $SO(3)$ based on direct inertial measurements," *Proc. of the IEEE International Conference on Robotics and Automation, Orlando, FL, USA, May 15-19, 2006, pp. 2170–2175*.

[48] D. Simon, *Optimal state estimation: Kalman, H_∞, and nonlinear approaches*. John Wiley & Sons, 2006.

[49] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[50] F. L. Markley, "Attitude error representations for Kalman filtering," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 26, no. 2, pp. 311–317, 2003.

[51] J. L. Crassidis and F. L. Markley, "Unscented filtering for spacecraft attitude estimation," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 26, no. 4, pp. 536–542, 2003.

[52] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," *Proc. of the American Control Conference, Seattle, WA, USA, June 21-23, 1995, pp. 1628–1632*.

[53] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. AC-45, pp. 477–482, March 2000.

[54] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter," *Kalman filtering and neural networks*, pp. 221–280, 2001.

[55] A. Doucet, *Sequential monte carlo methods*. Wiley Online Library, 2001.

[56] M. Sanjeev Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[57] Y. Cheng and J. L. Crassidis, "Particle filtering for sequential spacecraft attitude estimation," *Proc. of the AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, RI, USA, August 16-19, 2004*.

[58] Y. Oshman and A. Carmi, "Estimating attitude from vector observations using a genetic algorithm-embedded quaternion particle filter," *Proc. of the AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, RI, USA, August 16-19, 2004*.

[59] D. Firoozi and M. Namvar, "Analysis of gyro noise in non-linear attitude estimation using a single vector measurement," *IET Control Theory & Applications*, vol. 6, no. 14, pp. 2226–2234, 2012.

[60] M. Zamani, J. Trumpf, and R. Mahony, "Near-optimal deterministic filtering on the rotation group," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1411–1414, 2011.

[61] M. Zamani, J. Trumpf, and R. Mahony, "Minimum-energy filtering for attitude estimation," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2917–2921, 2013.

[62] K. J. Jensen, "Generalized nonlinear complementary attitude filter," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 34, no. 5, pp. 1588–1593, 2011.

[63] T. Lee, "Exponential stability of an attitude tracking control system on $SO(3)$ for large-angle rotation maneuvers," *Systems and Control Letters*, vol. 61, pp. 231–237, 2012.

[64] P. Tsiotras, "Further Passivity Results for the Attitude Control Problem," *IEEE Transactions on Automatic Control*, vol. 43, no. 11, pp. 1597–1600, 1998.

[65] J. C. Kinsey and L. L. Whitcomb, "Adaptive identification on the group of rigid-body rotations and its application to underwater vehicle navigation," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 124–136, 2007.

[66] P. C. Hughes, *Spacecraft attitude dynamics*. Mineola, New York: Dover, 2nd ed., 2004.

[67] R. N. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC Press, Inc., 1993.

[68] J. R. Wertz, *Spacecraft Attitude Determination and Control*. Dordrecht, The Netherlands: D. Reidel Publishing Company, 1978.

[69] S. P. Bhat and D. S. Bernstein, "A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon," *Systems and Control Letters*, vol. 39, pp. 63–70, 2000.

[70] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052–1067, 2007.

[71] M. Barczyk, S. Bonnabel, J. E. Deschaud, and F. Goulette, "Experimental implementation of an invariant extended Kalman filter-based scan matching SLAM," in *American Control Conference (ACC), 2014*, pp. 4121–4126, June 2014.

[72] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Transactions on Robotics*, vol. 33, pp. 1–21, Feb 2017.

[73] Y. Wang and G. S. Chirikjian, "Error propagation on the euclidean group with applications to manipulator kinematics," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, 2006.

[74] A. Barrau and S. Bonnabel, "An ekf-slam algorithm with consistency properties," *arXiv preprint arXiv:1510.06263*, 2015.

[75] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, pp. 1797–1812, April 2017.

[76] R. Mahony and T. Hamel, "A geometric nonlinear observer for simultaneous localisation and mapping," in *IEEE Conference on Decision and Control (CDC)*, pp. 2408–2415, IEEE, 2017.

[77] K. W. Lee, W. S. Wijesoma, and I. G. Javier, "On the observability and observability analysis of SLAM," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3569–3574, IEEE, 2006.

[78] L. W. Tu, *An Introduction to Manifolds*. Springer, 2011.

[79] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*, vol. 2. Springer Science & Business Media, 2011.

[80] R. Lozano, B. Brogliato, O. Egeland, and B. Maschke, *Dissipative systems analysis and control: theory and applications*. Springer Science & Business Media, 2013.

[81] J. P. La Salle, "An invariance principle in the theory of stability," in *Control theory, Twenty-Five Seminal Papers (1932-1981)* (T. Basar, ed.), pp. 309–320, New York: Wiley-IEEE Press, 2001.

[82] T. D. Barfoot, "State estimation for aerospace vehicles – AER1513 course notes and experimental datasets," *University of Toronto Institute for Aerospace Studies,* AER1513-002, Rev: 1.0, 2009.

[83] A. Barrau and S. Bonnabel, "Intrinsic filtering on lie groups with applications to attitude estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, 2015.

[84] A. Khosravian, T. J. Chin, I. Reid, and R. Mahony, "A discrete-time attitude observer on $SO(3)$ for vision and gps fusion," pp. 5688–5695, May 2017.

[85] A. H. de Ruiter and J. R. Forbes, "On the solution of Wahbas problem on SO(n)," *The Journal of the Astronautical Sciences*, vol. 60, no. 1, pp. 1–31, 2013.

[86] S. A. Holmes, G. Klein, and D. W. Murray, "An $O(N^2)$ square root unscented Kalman filter for visual simultaneous localization and mapping," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 7, pp. 1251–1263, 2009.

[87] M. Izadi and A. K. Sanyal, "Rigid body attitude estimation based on the Lagrange–d'Alembert principle," *Automatica*, vol. 50, no. 10, pp. 2570–2577, 2014.

[88] M. Izadi, A. K. Sanyal, E. Samiei, and S. P. Viswanathan, "Discrete-time rigid body attitude state estimation based on the discrete Lagrange-d'Alembert principle," in *American Control Conference (ACC), 2015*, pp. 3392–3397, IEEE, 2015.

[89] S. Berkane, A. Abdessameud, and A. Tayebi, "Hybrid global exponential stabilization on SO(3)," *Automatica*, vol. 81, pp. 279 – 285, 2017.

[90] W. T. Higgins, "A comparison of complementary and kalman filtering," *IEEE Transactions on Aerospace and Electronic Systems*, no. 3, pp. 321–325, 1975.

[91] M. Zimmermann and W. Sulzer, "High bandwidth orientation measurement and control based on complementary filtering," in *Robot Control 1991*, IFAC Symposia Series, pp. 525 – 530, Amsterdam: Pergamon, 1992.

[92] G. Tao and P. Ioannou, "Strictly positive real matrices and the lefschetz-kalman-yakubovich lemma," *IEEE Transactions on Automatic Control*, vol. 33, no. 12, pp. 1183–1185, 1988.

[93] J. T. Wen, "Time domain and frequency domain conditions for strict positive realness," *IEEE Transactions on Automatic Control*, vol. 33, pp. 988–992, October 1988.

[94] H. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.

[95] H. Marquez, *Nonlinear Control Systems*. Hoboken, NJ: John Wiley and Sons, Inc., 2003.

[96] A. Sanyal, A. Fosbury, N. Chaturvedi, and D. Bernstein, "Inertia-free spacecraft attitude tracking with disturbance rejection and almost global stabilization," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 32, no. 4, pp. 1167–1178, 2009.

[97] P. Pisu and A. Serrani, "Attitude tracking with adaptive rejection of rate gyro disturbances," in *American Control Conference, 2008*, pp. 4839–4844, IEEE, 2008.

[98] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*. CRC press, 2011.

[99] J. B. Hoagg, M. A. Santillo, and D. S. Bernstein, "Internal model control in the shift and delta domains," *IEEE Transactions on Automatic Control*, vol. 53, no. 4, pp. 1066–1072, 2008.

[100] K. Zhou and J. C. Doyle, *Essentials of robust control*, vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[101] S. Skogestad and I. Postlethwaite, *Multivariable feedback control: analysis and design*, vol. 2. Wiley New York, 2007.

[102] A. Saccon, J. Hauser, and A. P. Aguiar, "Optimal control on lie groups: The projection operator approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2230–2245, 2013.

[103] K. Zhou, J. C. Doyle, K. Glover, *et al.*, *Robust and optimal control*, vol. 40. Prentice hall New Jersey, 1996.

[104] R. Lozano-Leal and S. Joshi, "On the design of the dissipative LQG-type controllers," in *Proceedings of the 27th IEEE Conference on Decision and Control*, pp. 1645–1646, IEEE, 1988.

[105] W. M. Haddad, D. S. Bernstein, and Y. W. Wang, "Dissipative $\mathcal{H}_2/\mathcal{H}_\infty$ controller synthesis," *IEEE Transactions on Automatic Control*, vol. 39, no. 4, pp. 827–831, 1994.

[106] J. Geromel and P. Gapski, "Synthesis of positive real $\mathcal{H}_2$ controllers," *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 988–992, 1997.

[107] T. Shimomura and S. P. Pullen, "Strictly positive real $\mathcal{H}_2$ controller synthesis via iterative algorithms for convex optimization," *Journal of Guidance, Control, and Dynamics*, vol. 25, no. 6, pp. 1003–1011, 2002.

[108] C. J. Damaren, "Optimal strictly positive real controllers using direct optimization," *Journal of the Franklin Institute*, vol. 343, no. 3, pp. 271–278, 2006.

[109] J. R. Forbes and C. J. Damaren, "Design of gain-scheduled strictly positive real controllers using numerical optimization for flexible robotic systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 132, no. 3, p. 034503, 2010.

[110] J. R. Forbes and C. J. Damaren, "Design of optimal strictly positive real controllers using numerical optimization for the control of flexible robotic systems," *Journal of the Franklin Institute*, vol. 348, no. 8, pp. 2191–2215, 2011.

[111] J. R. Forbes, "Dual approaches to strictly positive real controller synthesis with a h2 performance using linear matrix inequalities," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 8, pp. 903–918, 2013.

[112] J. R. Forbes and C. J. Damaren, "Synthesis of optimal finite-frequency controllers able to accommodate passivity violations," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1808–1819, 2013.

[113] A. K. Sanyal, T. Lee, M. Leok, and N. H. McClamroch, "Global optimal attitude estimation using uncertainty ellipsoids," *Systems & Control Letters*, vol. 57, no. 3, pp. 236–245, 2008.

[114] L. J. Bridgeman and J. R. Forbes, "Conic-sector-based control to circumvent passivity violations," *International Journal of Control*, vol. 87, no. 8, pp. 1467–1477, 2014.

[115] L. J. Bridgeman, R. J. Caverly, and J. R. Forbes, "Conic-sector-based controller synthesis: Theory and experiments," in *American Control Conference (ACC)*, pp. 4292–4297, IEEE, 2014.